22 March 2001 | www.nature.com/nature | £3

nature



HIGH-SPEED POWER
Faster lifters: butterflies

FEED THE WORLD
The local solution

YOU'VE GOT RHYTHM
When body clocks go wrong

PEKING MAN WAS COOL

Older and 'colder' dates for classic *Homo erectus* fossils

SUSTAINLAND
Forest or plantation

nature

www.nature.com/nature

Vol 458 | Issue no. 7235 | 12 March 2009

By common consent

Engaging with the public helped US scientists build the consensus that finally overturned federal restrictions on human embryonic stem-cell research. That public outreach should not stop now.

President Barack Obama's new policy on stem-cell research and his strong statement on restoring integrity to science (see page 130) are victories for science that hold key lessons for future science-policy debates. But the battle over stem-cell research will not end here, and scientists must continue to engage the public to maintain the support that eventually convinced politicians to back this work.

President George W. Bush satisfied no one when he declared that US federal funding could be used to study only human embryonic stem-cell lines created before 9 August 2001. Scientists thought the cells held enormous potential to unravel the mechanisms of disease, and were frustrated at the arbitrary restriction, because it limited funding to 21 lines made with old techniques. The policy also forced scientists to waste money building facilities to separate research funded by the National Institutes of Health (NIH) from that funded privately. And although induced pluripotent stem cells (iPS cells) were created from adult cells in 2006, in part to sidestep the restrictions, they have yet to show all the same abilities as embryonic stem cells.

Those opposed to embryonic stem-cell research because it involves destruction of embryos were not satisfied with Bush's approach either — it still allowed scientists to use federal money to study cell lines whose derivation had involved harm to embryos.

Scientists correctly realized that they would not be able to overturn the policy without public support. Research advocacy groups linked with patient advocacy groups to form organizations such as the Coalition for the Advancement of Medical Research in Washington DC, which brought people suffering from diseases to Capitol Hill to argue in favour of the research. Testimony from people such as the late actor Christopher Reeve, and the advocacy of others, such as Nancy Reagan, whose husband, former President Ronald Reagan, suffered from Alzheimer's disease, were compelling to lawmakers. Realizing that most Americans hold moderate views in the abortion debate, scientists appealed to their practical nature, arguing that the NIH would fund research only on lines made from embryos that were going to be discarded from fertility clinics anyway.

These strategic moves gradually earned public support for the work, culminating in Congress twice passing legislation that would allow the NIH to fund the research. Although Bush vetoed these laws, it became almost inevitable that his arbitrary policy would one day be overturned.

Both sides in the debate were charged with hype throughout this process. Scientists were accused of falsely promising cures they could not guarantee would materialize. Those opposed to embryonic stem-cell research were attacked for exaggerating and misinterpreting reports about the power of adult stem cells. No one knows whether embryonic stem cells will yield cures. But the financial weight of the NIH and the support of Obama's administration will dispel a cloud that discouraged investors from backing the expensive development process necessary to find out.

"Scientists must continue listening to public concerns on research."

There are still formidable challenges ahead. Congress needs to be convinced to pass legislation cementing Obama's policy so that future presidents cannot restrict scientific research with the stroke of a pen.

Then there are issues with the work itself. For instance, regulation governing the research in different countries is still inconsistent, limiting scientific collaboration. Major questions remain about the feasibility of using the cells as therapies — many scientists now believe embryonic stem cells will be most useful in drug screening and disease modelling in the lab. But working out how to replicate disease processes that take place over a lifetime will require years of study. And although iPS cells seem highly promising, there is much work to be done to remove traces of the reprogramming factors used to create them, and to test how similar they are to embryonic stem cells.

The good news is that NIH-funded research on these questions is no longer limited by an arbitrary policy. And if scientists continue the winning strategy of listening to public concerns on research — and responding appropriately — they will maintain the public trust that will let them pursue these questions unfettered. ■

Smart thinking

The US electricity grid needs to evolve and requires fresh standards of communication.

This week, the US Senate began to craft a massive energy bill that would establish a cap-and-trade system for carbon emissions and would pour billions of dollars into renewable-energy projects and upgrades to the nation's electricity grid.

Those investments, which will come on top of the billions of dollars already provided in the recent stimulus bill, are long overdue. Lawmakers working on the new bill should pay close attention to countries such as Denmark and Germany, where the transition to energy systems based on renewable sources and small-scale generation is already far advanced.

Among the most important lessons from these countries is that the transition demands a substantial decentralization of the electric power grid. Instead of being organized around a comparatively small number of very big power plants, as it currently is in the United States and many other countries, the new grid will have to accommodate a much larger number of local sources. These include not just

© 2009 Macmillan Publishers Limited. All rights reserved

renewable sources such as wind and solar, but small-scale combined-heat-and-power plants, and electricity generated from the recovery of waste industrial heat (see page 138).

This decentralization, in turn, requires a grid that is not just bigger, but smarter. With so many sources and sinks for power — and with buyers, sellers and even programmed refrigerators constantly trying to optimize their energy usage based on the power's moment-to-moment price — the electricity-moving system also has to be an information-moving system. Everything from the smart appliances, to the digital meters that will track electrons entering and exiting the wires, to the software that gives operators an overview of energy flow, have to be able to communicate with each other.

And therein lies one of the big challenges: there are no generally accepted standards for such communication. Utilities need those standards before they can make serious investments in the smart grid, as they would otherwise risk throwing money away on obsolete technology. And the manufacturers of smart-grid components want those standards to be international, as they would rather not make a different version of each product for each country.

In the United States, responsibility for forging a consensus on how the components of the smart grid will talk to each other falls to the National Institute of Standards and Technology (NIST), which received $10 million for the task in the stimulus bill. NIST should move quickly. Every day that a coal-powered factory can't work out how to sell its excess heat back to the grid means more fossil-fuel energy is wasted. And in the absence of the improvements in distribution efficiency that the smart grid can provide, impatient advocates of renewables — and impatient politicians eager to provide jobs during a downturn — are likely to go overboard on building new transmission lines.

"Utilities need communication standards before they can make serious investments in the smart grid."

At least some new transmission lines are inevitable — the most productive spots for large-scale wind or solar installations tend to be a long way from the population centres where power is consumed. But it is also true that large-scale transmission can sometimes be traded off against small-scale distribution. Increasing the diversity and number of generators while using digital information to distribute electricity more efficiently may extend the usefulness of current wiring.

Add to that a smaller number of carefully sited new lines, and the result may be a grid to carry the world forward into a greener age. But policy-makers shouldn't forget: bigger is not necessarily smarter. ■

Delicate balance

China's autocratic approach to environmental issues may not always reap the desired rewards.

In an ironic twist, the decidedly authoritarian government of China has received much praise of late from Western environmentalists, who tend to cast themselves as anti-authoritarian types.

China has indeed made laudable efforts in many fields, ranging from wind energy to vehicle-emission standards and pollution control. It has even outlawed free plastic bags. The United Nations Environment Programme gave it rave reviews for its pollution reduction and green measures at the Beijing Olympics (see http://tinyurl.com/ckllzm). And now, China is planning to push wetlands conservation and restoration projects with a speed and lavishness that is making Western environmentalists envious (see page 134).

But any tally of China's environmental achievements requires a reality check. In some areas, for example, it is not clear whether China has been as successful as it claims. The government never really answered criticism that its much touted pre-Olympic increase in the number of relatively smog-free 'blue sky days' was achieved not by making the skies less polluted, but by moving the monitors to cleaner areas. Moreover, China tends not to listen to critics, which stops some of its well-intentioned projects from achieving their full potential. China could undoubtedly improve the efficiency and the effectiveness of its environmental projects if it opened itself up to more feedback, especially from scientists. A case in point is the nation's wind-energy initiative, which has been hampered by sub-standard turbine technology and poorly planned grid connections (see *Nature* **457,** 372–374; 2009).

Similar caution needs to be exercised when evaluating China's wetlands projects. China's government is full of engineering prowess, and it can design projects that are beautiful and often breathtaking in concept. But once it has committed itself, other considerations fade — including some that shouldn't. It is worth remembering that neither the 1.24 million people displaced over the past decade by the Three Gorges Dam, nor the ecologists concerned about the dam's environmental toll, had much of a voice in the decision-making. If the government takes a similarly reckless approach to water diversion for wetlands restoration, the project could once again tarnish any environmental good with social harm.

But even in the West, a similarly delicate balance has to be struck. The United States, for example, has been anything but decisive in its efforts to repair the coast of Louisiana, where the bayous were badly damaged by Hurricane Katrina in 2005. That is because doing so would require the forcible seizure of land from thousands of private owners for the construction of weirs and other water-diversion systems. So one value — the common good, which in this case means saving the coast and protecting the millions of people in New Orleans — is in direct conflict with the values of property rights and individual autonomy.

The challenge for environmentalists is to find creative mechanisms to reconcile those values, while respecting all of them. The cap-and-trade approach now being embraced for controlling carbon emissions is a prime example. It uses the market system to achieve pollution reduction while allowing for flexibility and individual choice. The 'environmental services' idea of trying to put an economic value on the benefits provided by, say, coastal wetlands is an approach that is similar in spirit, if less mature. Such mechanisms hold the promise of minimizing or eliminating the need for authoritarian methods — East or West. ■

© 2009 Macmillan Publishers Limited. All rights reserved

RESEARCH HIGHLIGHTS

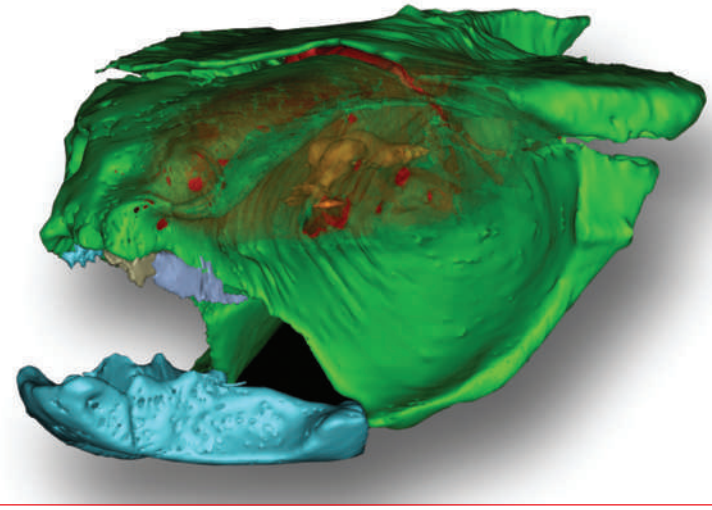
Brain box

Proc. Natl Acad. Sci. USA doi: 10.1073/pnas.0807047106 (2009)

A remarkably well-preserved brain has been discovered in a 300-million-year-old fossil of a fish from Kansas.

Philippe Janvier of the National Museum of Natural History in Paris and his colleagues used X-rays from the European Synchrotron Radiation Facility in Grenoble, France, to peer inside the skulls of iniopterygians, extinct relations of modern sharks and ratfish.

In one skull (pictured), they imaged a dense calcium phosphate structure. Its shape and relationship to nerve locations suggest that it is a brain, mineralized by phosphate-fixing microbes before the soft tissue could decay. The researchers hope that other vertebrates fossilized in similar conditions might yield further preserved organs, potentially throwing light on brain evolution.



PROC. NATL ACAD. SCI. USA/P. JANVIER/CNRS

CHEMICAL BIOLOGY

Sweet disguise

Nature Chem. Biol. doi:10.1038/nchembio.151 (2009)

An analysis of the sugars that cloak HIV-1 indicates that the virus leaves cells by usurping a native pathway used to spit out bits of cell membrane. This may help to camouflage it from the immune system.

When HIV-1 exits a cell, it picks up a coat of proteins decorated with sugars. Lara Mahal of the University of Texas at Austin and her colleagues found that the sugar profile of HIV-1 particles in cell culture matched that of microvesicles — cell membrane fragments thought to modulate immune responses — that are shed by infected cells.

The results raise concern that therapies aiming to block the interaction between HIV sugars and host cells could also interfere with microvesicle function. However, the researchers add that the same process may not be hijacked in all infected cells; they looked only at immune cells called T cells.

PROTEOMICS

Worm versus fly

PLoS Biol. **7,** e1000048 (2009)

The nematode worm *Caenorhabditis elegans* and the fruitfly *Drosophila melanogaster* (pictured, right) produce similar relative amounts of analogous proteins, even though levels of the messenger RNAs that code for these proteins vary widely between the species.

Michael Hengartner at the University of Zurich in Switzerland and his colleagues used mass spectrometry to analyse almost 11,000 *C. elegans* proteins (roughly half of the worm's predicted gene products). Of these, they selected nearly 2,700 for which mRNA data were available and compared their abundances with those of related proteins in the fruitfly. Protein ratios in the two model organisms correlated highly despite the two species' 600 million years of separate evolution, illustrating that regulating protein abundance is more important than maintaining gene-expression levels.

CHEMISTRY

Sprouting tubes

Nature Chem. doi:10.1038/nchem.113 (2009)

Tiny tubes grow spontaneously when small crystals of an inorganic solid are dunked in a solution of an organic ion. The effect could be exploited to build networks of miniature pipes that direct and control fluids.

Leroy Cronin and his colleagues at the University of Glasgow, UK, filmed tubes sprouting at up to 13 micrometres a second; some reached lengths of several centimetres.

The tubes grew from polyoxometalate (POM) crystals in solutions of phenanthridinium-based ions. The dissolving crystals first mingle with the ions to form a membrane, which subsequently bursts, squirting out POM that forms tube walls on contact with more organic ions.

The tubes can be bent with an electric field or the judicious use of obstacles. Their diameters can be customized by altering the concentration of organic ions.

MARINE ECOLOGY

Deadly dusting

Proc. Natl Acad. Sci. USA doi.10.1073/pnas.0811486106 (2009)

Atmospheric aerosols, which supply ocean phytoplankton with nutrients such as nitrates and iron, can have toxic effects too.

Adina Paytan at the University of California, Santa Cruz, and her colleagues found that some species of phytoplankton in Red Sea surface waters suffered when exposed to aerosol samples collected from the Sahara Desert in Africa, whereas they thrived on European aerosols. Copper may be to blame: it was present at much higher levels in African aerosols and can be toxic to phytoplankton.

Although desert dust supplies about two-thirds of atmospheric copper, human industrial emissions of the element are increasing rapidly. The researchers suggest that copper deposition could alter marine ecosystems by damaging phytoplankton in high-aerosol areas such as those downwind of industrial regions of south and east Asia.



MPG/EPA/CORBIS/D. M. PHILLIPS/VISUALS UNLIMITED/GETTY

OPTICS

Beyond the invisibility cloak

Phys. Rev. Lett. **102,** 093901 (2009)

Invisibility shields that cause objects within them to vanish are now well-known. But Yun Lai and his colleagues of the Hong Kong University of Science and

© 2009 Macmillan Publishers Limited. All rights reserved

Technology in Clear Water Bay show how such a device might work to hide an object outside its confines.

In their theoretical scheme, the custom-made shield creates an 'anti-object' image that cancels out the light-scattering of the hidden object nearby, so that in optical terms the whole system is replaced by empty space.

The researchers provide a prescription for making anti-objects of arbitrary shape embedded in slabs of metamaterials, which are composed of building blocks that interact with light in unusual, tailor-made ways.

PALAEONTOLOGY

Bird in the hand

PLoS ONE **4,** e4591 (2009)

Prints left by a squatting bipedal dinosaur show clearly that the hands faced inwards, as birds' limbs do now to allow for wing folding. The prints, which are almost 200 million years old, are preserved in Utah sandstone.

The dinosaur was a theropod — one of a group of mainly carnivorous dinosaurs, such as tyrannosaurs and velociraptors, that were the ancestors of modern birds. The prints suggest that theropods exhibited bird-like anatomy and resting postures much earlier than previously recognized.

Andrew Milner, curator at Utah's St George Dinosaur Discovery Site, and his colleagues theorize that the unknown creature's hand orientation was for grabbing or holding prey.

NEUROBIOLOGY

Second fiddle

Science **323,** 1313–1319 (2009)

Ion channels known as AMPA receptors help to transmit fast excitatory nerve impulses in the brain. Changes to their properties, which are regulated by other proteins, are crucial to many processes, including some involved in learning and memory.

TARP proteins were thought to be the only candidates for AMPA-receptor regulation. Yet in the rat brain such proteins associate with only about 30% of AMPA receptors.

Bernd Fakler and Nikolaj Klöcker at the University of Freiburg in Germany and their colleagues have unexpectedly come across what turns out to be the predominant partner proteins, using a quantitative proteomics approach. They found that, in rat brains, about 70% of AMPA receptors associate with cornichon proteins. Cornichons and TARPs regulate the function and expression of AMPA receptors differently, potentially allowing fine-tuning of fast nerve-impulse transmission.

ECOLOGY

Open goal

Proc. R. Soc. B doi:10.1098/rspb.2008.1762 (2009)

A host of grasses, including crops such as maize (corn) and sugarcane, grow well in hot, dry conditions because they have evolved a more efficient type of photosynthesis, the C_4 pathway, which requires less water than the more widespread C_3 pathway.

Some scientists believe that C_4 photosynthesis evolved to cope with drought, but others argue it was originally an adaptation to exposed habitats.

Colin Osborne and Robert Freckleton at the University of Sheffield, UK, have now settled the debate by analysing the habitats of 117 genera of grasses, representing 15 independent C_4 lineages. They provide strong evidence that C_4 plants first arose in open, tree-less environments. Once the pathway had evolved, however, C_4 plants would have been able to adapt faster to dry environments than C_3 plants.



N. CATTLIN/FLPA

MATERIALS SCIENCE

Diaphite domains

Phys. Rev. Lett. **102,** 087402 (2009);
Phys. Rev. B **79,** 054111 (2009)

Graphite, a form of carbon made up of layered sheets of atoms, transforms into diamond — another form — only when subjected to high pressure. But Katsumi Tanimura at Osaka University in Japan, Hiromasa Ohnishi at the Institute of Materials Structure Science in Tsukuba and their colleagues have found that by firing femtosecond (10^{-15} s) laser pulses at a graphite wafer, they can cause five-nanometre-diameter patches of the flat layers to buckle and become three-dimensional.

The patches, which form a structure that exhibits bonding between layers but differs from a diamond lattice, have been termed 'diaphite' domains. The material is stable at room temperature if kept under ultra-high vacuum.

JOURNAL CLUB

Frank Wilczek
Massachusetts Institute of Technology, Cambridge

A theoretical physicist examines exotic particles lurking in new materials.

Axions are very light, very weakly interacting particles, whose existence was posited more than 30 years ago in order to clean up our 'standard model' of particle physics. They close an annoying loophole in Kobayashi and Maskawa's Nobel prizewinning explanation of why the microscopic laws of physics look so nearly the same when running backwards as forwards in time (time reversal symmetry).

Despite heroic efforts — and several false alarms — axions have not yet been detected, but they have become increasingly important. They have been warmly embraced in unified field theories and in string theory. And when we run the equations through Big-Bang cosmology, we find that axions should contribute much of the dark matter that astronomers have inferred to explain the Universe.

Now Shou-Cheng Zhang and his colleagues (X.-L. Qi *et al. Phys. Rev. B* **78,** 195424; 2008) inform us that, all along, axions have been lurking unrecognized on surfaces of bismuth–tin alloys and other materials. To be more precise: the equations that arise in axion physics are the same as those that describe the electromagnetic behaviour of a recently discovered class of materials known, collectively, as topological insulators.

The axion field inside topological insulators is an emergent — and subtle — property of collections of electrons that is connected to their spin–orbit coupling.

These 'quasi-axions' don't improve our standard model, but they do have the charming advantage of being accessible, possibly even useful. There are ideas to exploit their behaviour to make anyons, potential building blocks for quantum computation.

No short summary can do justice to the wealth of ideas synthesized in this paper. Powerful, beautiful mathematics is at play in reality.

Discuss this paper at **http://blogs.nature.com/nature/journalclub**

© 2009 Macmillan Publishers Limited. All rights reserved

NEWS



Heather Melichar (left) and Ou Li at the University of California, Berkeley, can now use the federally funded equipment in their main lab to do embryonic stem-cell research.

P. SKORPINSKI

Obama overturns stem-cell ban

President's executive order will allow US human embryonic stem-cell research to thrive at last.

Scientists and research advocates worldwide are celebrating the removal of rules limiting research on human embryonic stem cells in the United States, which they say have restricted the field's progress for seven and a half years.

The National Institutes of Health (NIH) in Bethesda, Maryland, is now working out policies that will allow researchers to apply for grant money from the agency to study some of the hundreds of cell lines created since 9 August 2001, when President George W. Bush limited federal funding to research on lines in existence at that time. Some scientists are already proposing to use the new lines in applications for $200 million in NIH 'Challenge' grants, which will be funded by the economic stimulus package signed into law last month. Details of these grants were unveiled last week (see http://tinyurl.com/9bhdck).

Estimates of the number of new lines range from 400 to 1,000. Unlike the 21 lines previously eligible for federal funding, many of the lines have been made from embryos that had genetic predispositions to specific diseases, or were derived using 'animal-free' preparations, and thus could be more relevant to laboratory research and preclinical studies.

President Barack Obama signed the executive order on 9 March at a White House ceremony attended by scientists, lawmakers, patients and patient advocates. "We will vigorously support scientists who pursue this research," Obama said. "And we will aim for America to lead the world in the discoveries it one day may yield."

The new order asks the NIH to develop guidelines and regulations within 120 days to govern federally funded human embryonic stem-cell research. Work is already under way at the NIH to develop guidelines covering the eligibility of cell lines for federal funding. These will be based on issues such as the kind of informed consent given by couples who donated the leftover embryos from which the cells were collected. Such cells can develop into any type of tissue in the body, and are thus thought to hold enormous promise as tools for dissecting disease processes, screening possible treatments and developing new therapies.

Obama order deals with scientific integrity

On the day that President Barack Obama lifted restrictions on the federal support of human embryonic stem-cell research, he also issued a memo directing the head of the Office of Science and Technology Policy — the position for which John Holdren has still to be confirmed — to ensure scientific integrity in government decision-making.

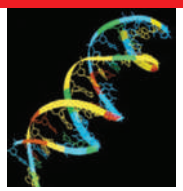
The memo calls for agencies to hire scientific employees based on their knowledge, and to put rules in place for scientific integrity and transparency along with protection for whistleblowers.

Harold Varmus, co-chair of the President's Council of Advisors on Science and Technology, says the memo would help prevent the kind of instances during the administration of George W. Bush when political appointees interfered with the workings of government scientists. "The intent is to prevent that kind of undermining of the science advisory process," he says.

But John Marburger, Bush's science adviser, said in an e-mail: "There's certainly nothing wrong with it, but I never believed such a memorandum was necessary. From the perspective of the Obama administration, it delivers on a campaign commitment."

Robert George, a bioethicist at Princeton University in New Jersey and a member of the President's Council on Bioethics, says that few would complain about getting the best scientific advice or maximizing transparency. But he says that the memo could run the risk of tarring opponents of stem-cell research as ideologues and enemies of science. "It's not a question of science on one side and ideology on the other," he says. "It's a dispute about what ethical norms will govern science." **E.C.H.**

© 2009 Macmillan Publishers Limited. All rights reserved



GENETIC RESURRECTION
Lost DNA gets a second chance.
www.nature.com/news

PUNCHSTOCK

Legislation to codify the change has already been introduced into both the House of Representatives and the Senate. It explicitly permits federal funding for research on stem-cell lines derived with parental permission from embryos left over at fertility clinics and otherwise slated for destruction. At least one observer has suggested that legislation explicitly approving federal funding for stem-cell research is needed to address the Dickey–Wicker amendment, a law first enacted by Congress in 1996 and renewed every year since, that prohibits federal funding of research in which embryos are created or destroyed (see *Nature* **457,** 1068–1069; 2009).

Scientists and advocates say the change will speed up research on stem cells and the development of possible treatments around the world, and will help remedy what they see as the Bush administration's legacy of political interference in science. "This is the first time that the key decisions and guidelines on this issue are going to be made in Bethesda and not at 1600 Pennsylvania Avenue," says Larry Soler, executive vice-president of the Juvenile Diabetes Research Foundation International, based in New York City. "That is what the scientific community and the patient community have been asking for."

But there was a bittersweet taste to the victory, as scientists lamented almost a decade of delays to the research. "There's no doubt that this federal policy has really slowed the rate of progress of stem-cell research in the United States, and the impact of the change in policy would have been greater if it had happened years ago," says Arnold Kriegstein, director of the Eli and Edythe Broad Center of Regeneration Medicine and Stem Cell Research at the University of California, San Francisco.

Those who oppose the research because it involves the destruction of embryos criticized Obama's decision. They say that the NIH should support research only on cells that are not derived from embryos, such as induced pluripotent stem cells (iPS cells), which are derived from adult cells but have many properties of embryonic stem cells.

But iPS cells, first created in 2006, are not a substitute for embryonic stem cells, Kriegstein says. "iPS technology and its ongoing improvement will likely eclipse embryonic stem-cell lines for diagnostic and therapeutic applications, but for now, embryonic stem cells are clearly needed. And it's still not clear how iPS cells will ultimately compare for therapeutic purposes."

Obama's action will begin to level the playing field for researchers in the United States, where some states — most notably California, which created the $3-billion California Institute for Regenerative Medicine (CIRM) in 2004 — have leapfrogged ahead of others by funding stem-cell research themselves. And it will have ripple effects around the world, some predict. "This type of science is international, and the whole world has suffered from the previous short-sighted and rather bizarre policy," says Robin Lovell-Badge of the National Institute for Medical Research in London.

For instance, Lovell-Badge collaborates on one project with NIH-funded scientists who have had to limit their work to the previously approved cell lines. More broadly, the lack of policy guidance from the NIH — normally a global science-policy leader — "has contributed to the morass of different rules adopted around the world, which in turn inhibits collaboration, goodwill and progress", Lovell-Badge adds.

Indeed, CIRM president Alan Trounson says that other nations with restrictive stem-cell research policies, such as Germany, now seem to be rethinking their stance, and he predicts that the Obama administration's support for the research will accelerate this trend. Such support may also encourage wary investors and pharmaceutical companies to become more involved in funding the research, and this could benefit both the biotechnology industry and the development of new treatments.

"Since Obama was elected, the pharmaceutical industry is clearly much more interested in stem cells," Trounson says. "That will be a really big help when we're working through costly and difficult clinical trials to get treatments to patients." ■

Erika Check Hayden

See Editorial, page 125.

For more images of a US lab starting up stem-cell work again, see http://tinyurl.com/avleu8.

Personalized cancer therapy gets closer

The long-awaited era of personalized genetic medicine may finally be arriving for people with cancer. Some cancer centres are preparing to screen all patients for genetic glitches associated with the disease, and scientists are starting to use detailed information about patients' tumour genomes to decide which treatments might benefit them most. "Oncology is absolutely farther down this road of personalized medicine than other areas," says Joan Scott, deputy director of the Genetics & Public Policy Center in Washington DC.

Treatments that target specific mutant genes have been available for certain cancers since 2001, and a handful of studies have pinpointed mutations that reveal which tumours will respond best. For instance, mutations in the epidermal-growth-factor receptor (EGFR) gene affect how well patients with lung cancer respond to drugs that target the EGFR protein. But recent studies have highlighted the shortcomings of this one-gene-at-a-time approach by showing that the genetic roots of many cancers are both complicated and blurred: a wide variety of rare mutations can cause any one type of cancer[1], and mutations in the same genetic pathways occur in many tumour types.

Last week, the Massachusetts General Hospital in Boston announced that it will carry out broad genetic testing of almost all patients with cancer, screening for 110 mutations in 13 cancer-related genes (see table, overleaf). Other hospitals are also moving in this direction: the Memorial Sloan-Kettering Cancer Center in New York City, for example, already tests 40 mutations in seven genes in all patients with lung or colorectal cancers. By picking up mutations in key genetic pathways, the tests could help oncologists to choose drugs that have been approved or are in clinical development. "In the next few years, I think every major cancer centre is going to work on this approach," says molecular pathologist John Iafrate of Massachusetts General Hospital.

Such advances are possible because the cost of genetic screening and sequencing technologies has been dropping precipitously. Hospitals can amplify and read out single altered 'letters' — or point mutations — in the DNA sequences of tumour genes commonly implicated in certain cancers. And academic researchers are a step beyond this, using high-speed sequencing to collect information about every gene in individual patients' tumours.

Last month, for example, Marco Marra, director of the British Columbia Cancer Agency's genome sciences centre in Vancouver, told attendees at the Advances in Genome Biology and Technology meeting ▶

© 2009 Macmillan Publishers Limited. All rights reserved

on Marco Island, Florida, that his group had sequenced the whole genome of a tumour that had spread from one patient's mouth to his lung. The group had also used gene-expression studies to compare the activity of his tumour genes to those in healthy tissue.

They found that the patient had mutations in a tumour-suppressor gene called *PTEN*, and abnormally high expression of a gene downstream of *PTEN*, called *RET*. This explained why he had not responded to treatment with the drug erlotinib, and fits with some earlier studies suggesting that patients with active *PTEN* respond better to erlotinib[2]. Marra's team recommended instead that the patient be put on a drug called sunitinib, which inhibits the protein made by *RET*. The patient's cancer subsequently regressed.

"This work illustrates how one can potentially use next-generation sequencing technologies to establish the appropriate course of cancer treatment for individual patients," says Eric Green, scientific director of the National Human Genome Research Institute in Bethesda, Maryland.

Studies such as Marra's are being conducted in many patients to find genetic traits that could predict how tumours respond to treatment. And such studies are likely to increase as the cost of genome sequencing drops. Rick Wilson, director of the Genome Sequencing Center at Washington University in St Louis says that it cost more than $1 million for his centre to sequence the first tumour genome last year[3]. He estimates that the second tumour genome sequenced at the university cost $500,000, and projects that this could drop to $50,000 per tumour by early next year.



BURGER/PHANIE/REX FEATURES

Some patients with cancer have gene mutations that can be targeted by specific drugs.

Some businesses are finding that a handful of patients — who may already be paying thousands of dollars for their diagnosis and treatment — are willing to pay this sort of money for cancer-related genetic information. Last year, for example, a company called CollabRx, based in Palo Alto, California, began offering a service that analyses 15,000 genes, looking at gene expression, copy number and single-base mutations in individual patients' tumours for between $50,000 and $100,000. Five patients have used the service, which aims to determine which available drugs might be most suitable. Jay Tenenbaum, the company's founder and chairman, calls CollabRx's approach "personalized oncology research" because the results of each patient's individual analysis are used to help guide future analyses in other patients.

But because only a few drugs can currently be selected on the basis of specific mutations, "you have to be careful with raising expectations with that kind of approach", says Marc Ladanyi, chief of molecular diagnostics at the Memorial Sloan-Kettering Cancer Center.

Anna Barker, deputy director of the National Cancer Institute in Bethesda, is optimistic that treatment options won't stay limited for long. Genomic data being gathered from patients with cancer should help researchers to identify the genetic signatures that correlate with different symptoms or rates of disease progression — and hence guide the development of drugs. Barker predicts that these signatures will also be used in the future to help doctors to combine existing drugs to target multiple genetic pathways and attack specific attributes of each patient's cancer. "I'm very optimistic that we can move in this direction, and it should be a more cost-effective and better way of taking care of cancer patients," she says. ■

Erika Check Hayden

PATIENT CANCER GENES UNDER SCRUTINY

Gene tested (by whom)	Implicated in which major cancers
APC (MGH)	Brain, colorectal, liver, stomach, thyroid
Beta-catenin (MGH)	Brain, colorectal, liver, lung, ovarian, skin
BRAF (MGH, MSKCC)	Colorectal, lung, skin, thyroid
EGFR (MGH, MSKCC)	Lung
FLT3 (MGH, MSKCC)	Blood
JAK2 (MGH, MSKCC)	Blood
KIT (MGH)	Blood, gastrointestinal, germ cell
KRAS (MGH, MSKCC)	Colorectal, lung, pancreas
NOTCH1 (MGH)	Blood
NRAS (MGH)	Blood, colorectal, thyroid
TP53 (MGH)	Anal, bladder, bone, brain, breast, cervical, colorectal, head and neck, liver, lung, skin, soft tissue
PIK3CA (MGH, MSKCC)	Brain, breast, colorectal, liver, lung, ovarian, stomach
PTEN (MGH)	Blood, brain, breast, cervical, endometrial, head and neck, prostate, skin
MEK1 (MSKCC)	Lung
AKT1 (MSKCC)	Breast, colorectal, lung, ovarian
HER2 (MSKCC)	Breast, ovarian

MGH, Massachusetts General Hospital; MSKCC, Memorial Sloan-Kettering Cancer Center.

1. Check Hayden, E. *Nature* **455,** 148 (2008).
2. Mellinghoff, I. K. *et al. N. Engl. J. Med.* **353,** 2012 (2005).
3. Ley, T. J. *et al. Nature* **456,** 66–72 (2008).

© 2009 Macmillan Publishers Limited. All rights reserved

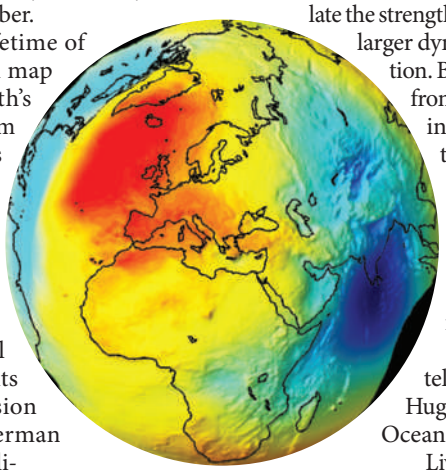
Gravity mission to launch

The planned launch next week of GOCE, a satellite designed to measure Earth's gravity field in unprecedented detail, will bolster the fields of geodesy, oceanography and climate prediction, scientists hope.

Part of the European Space Agency's Living Planet Programme, GOCE — for Gravity field and steady-state Ocean Circulation Explorer — is scheduled for launch on 16 March from the Plesetsk Cosmodrome in Russia, using a converted Russian SS-20 missile. Problems with its upper-stage navigation system delayed the launch from last September.

Over its expected lifetime of 20 months, GOCE will map tiny variations in Earth's gravity field that stem from the position of mountains and ocean trenches, and from small density variations in the planet's interior. Three pairs of cube-shaped accelerometers, in free fall inside the satellite, will make the measurements at five times the precision of GRACE, the US–German Gravity Recovery and Climate Experiment that has been orbiting Earth since 2002. The two missions are complementary; GOCE provides much better resolution at the small scale, whereas GRACE can show larger-scale trends, such as how gravity changes as ice sheets melt or gain mass.

ESA



The GOCE satellite will measure Earth's gravity field in unprecedented detail.

Permanent record

At a cost of €350 million (US$442 million), GOCE will determine the Earth's geoid — an idealized surface representing a level, motionless ocean — with an accuracy of 1–2 centimetres and at a spatial resolution of better than 100 kilometres. The data should enable scientists to map sea level better and to determine mean ocean circulation. "Until now, we could measure only the variable component of ocean currents such as the Gulf Stream, whereas the permanent part remained unknown," says Marine Herrmann, an oceanographer at the National Centre for Meteorological Research in Toulouse, France.

The data should also help to inform discoveries from other satellites, such as the Jason series that measures sea-level variation to deduce factors such as ocean convection.

With the geoid as a reference, scientists should be able to differentiate between signals due to climate change and those due to natural fluctuations in ocean circulation. "The ocean is the memory of the climate system," says Jochem Marotzke, an oceanographer at the Max Planck Institute for Meteorology in Hamburg, Germany. "Our goal is to make decadal climate prediction, but to do this you need to know the initial state of the ocean."

Computer models will still be needed to translate the strength of surface flows into the larger dynamics of ocean circulation. But much can be deduced from space. For instance, the intensity of deep convection — the process by which cold water sinks to depth — is related to sea surface elevation and can be monitored, at least approximately, from space[1].

"The surface itself tells you a lot," says Chris Hughes, from the Proudman Oceanographic Laboratory in Liverpool, UK, and a member of the GOCE mission advisory group. Earlier this year, Hughes showed that the sea level off the east coast of North America contains a signal of the strength of the Atlantic overturning circulation[2]. Gravity data will help to improve such estimates, he says.

GOCE's maximum lifetime is limited to two years, because it needs to fly at an altitude of 270 kilometres (compared with GRACE's initial 500 kilometres) to measure gravitational changes so precisely, and extra fuel is needed to compensate for the drag the atmosphere exerts at that low elevation.

"The ocean is the memory of the climate system."

Meanwhile, the successor to the current Jason-2 altimetry mission, Jason-3, remains some €60 million short of funding. The European Organisation for the Exploitation of Meteorological Satellites, one of the partners for Jason-3, plans to meet on 1 July to discuss how to divide up that shortfall between its member countries. ■

Quirin Schiermeier

1. Herrmann, M., Bouffard, J. & Béranger, K. *Geophys. Res. Lett.* **36,** L03606 (2009).
2. Bingham, R. J. & Hughes, C. W. *Geophys. Res. Lett.* **36,** L02603 (2009).

© 2009 Macmillan Publishers Limited. All rights reserved

Putting China's wetlands on the map

In his attempt to track avian flu in China, Peng Gong ended up creating the only comprehensive wetlands map the country has. When Gong compared that map, described in an article published in Chinese last month, to another as-yet unpublished one, he found that nearly 30% of China's natural wetlands vanished between 1990 and 2000.

The Chinese government plans to create its own, even more comprehensive, maps as it pumps money into wetlands restoration and conservation projects. Together, the new research will fill a hole for scientists and conservationists who had to watch as marsh and other wetlands were developed, often as farmland, without knowing the full extent of the loss.

Gong, an environmental scientist who holds a joint position at the University of California, Berkeley, and the Institute of Remote Sensing Applications in Beijing, took on the project because he wanted to track migrating birds, which are suspected of spreading avian influenza and which often spend summers or winters in wetlands.

But he had little to go on. A national wetland survey conducted between 1995 and 2003 merely listed total area in square kilometres attached to political districts. Other maps focused only on large lakes; for tracing avian flu, Gong needed more detail to see, for example, where there might have been smaller wetlands next to poultry farms. So he compiled Landsat satellite imagery of China taken around 1990 and 2000.

Demin Zhou, a hydroecologist at the Northeast Institute of Geography and Agroecology in Changchun, helped ground Gong's satellite imagery by providing detailed information on local water features. "It will be a goldmine for us and for other researchers," he says.

"Nearly 30% of China's natural wetlands vanished between 1990 and 2000."

For now, the gold is in documenting how wetlands have shrunk. In the Sanjiang Plain wetland in the northeast, where endangered red-crowned cranes spend their summers, Gong notes a drop from 22,932 km^2 to 10,114 km^2. "It's a huge threat," he says. Melting glaciers increased wetlands in regions such as Tibet and Qinghai, but the glacier retreat would also disturb the ecosystem (Z. Niu *et al. Science in China Ser. D: Earth Sci.* in press; 2009).

The map has its shortcomings. Given the limited field work that cross-checked satellite data, Gong could not categorize the wetlands in the same detail as given in the Ramsar Convention on Wetlands. A team led by Shuqing An, a wetland ecologist at Nanjing University, has found at least 26 natural and 9 man-made types of wetland in China. An says that Gong's high-resolution map will pick up a lot of small wetlands, but that their lack of detail will probably lead to misidentification of wetland types and make their use "confusing".

Gong says he would like to release the map in full detail (a lower resolution version is available at http://tinyurl.com/bbuszg) publicly later this year. But China is guarded about its geographical data (see *Nature* **451,** 871; 2008), and he needs to get approval from the State Bureau of Surveying and Mapping. He also plans to update the map with data from 2008.

The Chinese State Forestry Administration started a wetland mapping project last October, with an investment of 100 million renminbi (US$14.6 million). This will include remote-sensing data and much denser ground data than Gong's. The maps are meant to help organize restoration, conservation and 'wise use' initiatives, including ecological tourism and wetland national parks, for which the government has set aside 16.5 billion renminbi for its 2006–10 five-year plan.

Meanwhile, some Chinese institutions are moving quickly to set up international collaborations in wetlands research. One university — which did not want to be named until plans are finalized — is forming a new institute to span life, environmental, physical and social sciences, and policy studies, and hopes to collaborate with the Wetlands Biogeochemistry Institute at Louisiana State University in Baton Rouge.

A starting point for collaboration could be *Spartina alterniflora*, a salt-marsh grass introduced to China in the 1960s for land reclamation and whose spread has caused higher levels of sediment accumulation in the Yangtze delta. In its native habitat, along the US Mississippi River delta, the grass does not wreak such havoc. Irving Mendelssohn, a plant ecologist at Louisiana State University, says that comparative studies on topics such as *S. alterniflora* and nutrient cycling will be made before moving on to how best to conserve or restore wetlands. "The basic concept is to compare ecology, management and restoration of the coastal Yangtze River delta with the Mississippi River delta," he says. Last December, Tulane University in New Orleans and East China Normal University in Shanghai also signed a statement of intent for a collaboration comparing these two deltas.

The speed at which China seems ready to move has impressed foreigners. Douglas Meffert, an environmental scientist at Tulane, hopes their collaboration will accelerate the laggard 'master plan' for water-diversion projects to fix wetlands along the US Gulf Coast. Strategies tested in China could convince those in the United States to act. "China can move with much more alacrity to pull off large-scale restoration projects," he says. ■

David Cyranoski

See Editorial, page 126.



K. SU/CORBIS

Red-crowned cranes on their way to their summer home — or what's left of it.

© 2009 Macmillan Publishers Limited. All rights reserved



J. B. RUSSELL/PANOS

OPINION
Sharing the responsibility for China's carbon emissions.
www.nature.com/news

Web usage data outline map of knowledge

When users click from one page to another while looking through online scientific journals, they generate a chain of connections between things they think belong together. Now a billion such 'clickstream events' have been analysed by researchers to map these connections on a grand scale.

The work provides a fascinating snapshot of the web of interconnections between disciplines, which some data-mining experts believe reveals the degree to which work that is not often cited — including work in the social sciences and humanities — is widely consulted and can form bridges between scientific disciplines.

The creators of the maps argue that web-usage metrics give an alternative and more up-to-date view of science than existing maps and indicators, which are largely based on out-of-date citation data. Other researchers agree that the new maps, published this week (J. Bollen *et al. PLoS ONE* **4,** e4803; 2009), are impressive in approach, but they disagree on their significance.

For the study, Johan Bollen and his colleagues at the Los Alamos National Laboratory in New Mexico negotiated access to anonymized server log data covering 35,000 journals from 2006 to 2007. The data came from the University of Texas, the California State University system, and major science journal gateways including Thomson Reuters' Web of Science and Elsevier's Scopus database.

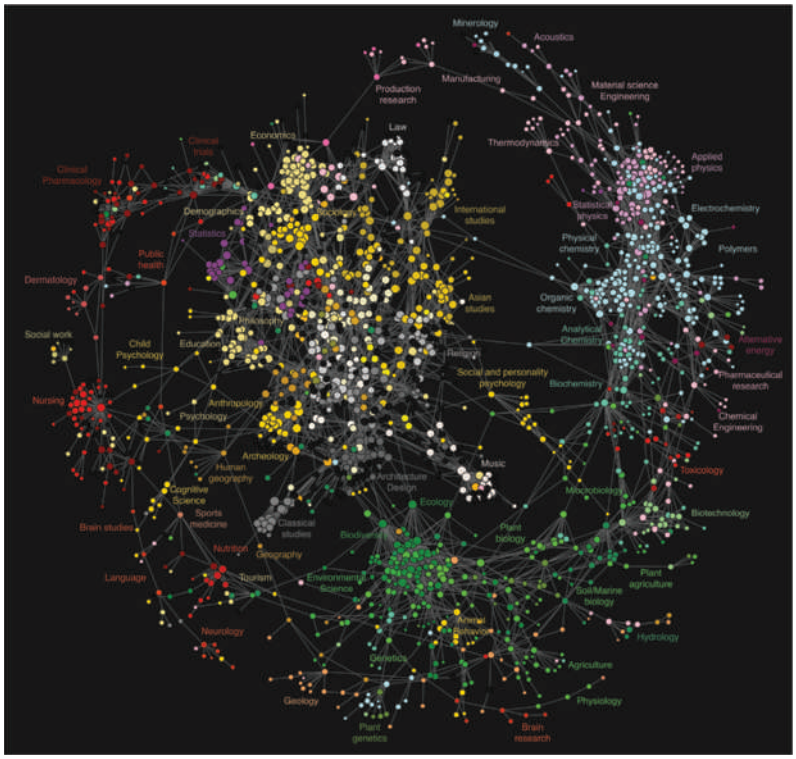
Although data on usage rather than citations have been used in some past studies, the sheer scale of the new study makes it stand out, says Henk Moed, a bibliometry expert at the Centre for Science and Technology Studies at the University of Leiden in the Netherlands. "The paper represents an important step forward."

The data reveal how often users looking at an article in journal A moved on to an article in journal B, and on to one in journal C, and so on, during a browser session. By aggregating hundreds of millions of such relationships, the researchers could use network-visualization algorithms to create maps based on the 'distances' computed between journals and disciplines.

The broad structure of the maps is similar to those created using citation data: a network of clusters in different fields, within which journals have strong connections with one another but fewer links to other clusters. A striking difference in the usage maps is that journals in the humanities and social sciences figure much more prominently than in citation-based maps. Along with some journals in other fields, such as psychology and the environment, they also emerge as gateways between clusters that are otherwise poorly connected, and so act as key bridges between disciplines. The difference partly arises because Bollen's study covers a wider literature than the citation databases, which are biased towards natural sciences journals.

The journal ranking generated from the usage maps includes not just the usual suspects such as *Nature*, *Science* and *Physical Review B*, but also the *Journal of Advanced Nursing* and *Environmental Health Perspectives*. That reflects a key difference between citation- and usage-based maps and metrics. The former reflect citations by researchers who publish, but ignore the impact of papers on large swathes of the scientific and medical community who read and apply the literature in medical, commercial or policy practice but who rarely or never publish.

"Citation data may undervalue papers written in practitioner-based fields, such as reviews or syntheses in clinical medical journals that are widely read by practising physicians but not cited proportionally," says Carl Bergstrom of the University of Washington in Seattle. "By including practitioners we capture a much wider sample of the scholarly community," adds Bollen.



Usage maps may give a truer picture of the information people access than citation does.

PLOS ONE

Usage maps are also more up to date than citation ones because the inherent delay in publication means it takes at least two years before a paper will start to gather citations in sufficient numbers to be meaningful. "The most exciting aspect is that they give us a different time-slice of the process of scientific discovery," says Bergstrom.

Others are less impressed. Anthony van Raan, director of the Leiden Centre for Science and Technology Studies, argues that this more current view may in fact represent today's "fashions", rather than trends that will endure. Faster online publishing means that papers are being cited faster than before, he argues. He also questions the central position of the social sciences in the maps, and various aspects of the data-analysis techniques used. Other experts say they have similar concerns, but are holding off from passing judgement until they can discuss the methodology with the paper's authors.

But Bergstrom argues that usage and citation data each provide different but useful information on the impact of papers and journals. "Usage data tell us where the net was cast; citation data tell us where the fish were caught," he says. "If you want to understand the human enterprise of fishing, you had better know about both." ■

Declan Butler

© 2009 Macmillan Publishers Limited. All rights reserved

Obama appoints first federal IT chief

The BlackBerry-toting Barack Obama last week took a step towards modernizing the US government's information system by appointing Vivek Kundra to the newly created post of chief information officer.

Kundra, who for the past two years was chief technology officer (CTO) for Washington DC, will have authority over IT spending across the entire federal government, a sum that currently comes to $71 billion per year. He will report to the director of the White House Office of Management and Budget.

"For those of us who have been hoping to see progress in Obama's stated vision of rebooting the federal bureaucracy, [Kundra's appointment] is like winning the World Series," says Andrew Rasiej, founder of Personal Democracy Forum, an annual conference and website on politics and technology.

Indeed, Kundra has won widespread praise for his innovations in Washington, which include getting city agencies to provide public data in standard formats such as XML and RSS feeds.

Kundra also embraced 'cloud computing' by moving all 38,000 employees for Washington DC onto the Google Apps office suite, which stores data and applications online rather than on individual computers. And he recently implemented an internal management tool in which information about the status and budget of each city project is displayed in a uniform, easy-to-read format.



Vivek Kundra is to be the US chief information officer.

GOVT OF DISTRICT OF COLUMBIA

"The one real measure of a CTO is their accomplishments," says Bob Gourley, CTO for the consulting firm Crucial Point based in Manassas, Virginia, and author of the blog CTOvision.com. "And Vivek has racked them up."

"There are three pillars to my agenda," Kundra told *Nature* in an interview last December. "Engaging citizens, driving transparency and greatly decreasing the cost of government operations." In a 5 March news conference announcing his appointment, he promised to take much the same approach at the federal level. For example, he says, "We want to launch a data.gov site to make a vast array of government data public."

However, Kundra also warned that changing a $71-billion operation will take time. Any federal reliance on cloud computing, for example, will have to be evaluated in the light of security requirements. Likewise, any data streams coming out of data.gov will have to comply with legislation that protects citizens' privacy. ■

M. Mitchell Waldrop

© 2009 Macmillan Publishers Limited. All rights reserved

Merck strengthens drug pipeline in rival takeover

Pharmaceutical giant Merck will acquire its rival Schering-Plough in a US$41-billion deal announced on 9 March and unanimously approved by the boards of the New Jersey-based companies.

The move is seen as a way to protect the company's drug-development pipeline. The new company, which will be called Merck, would have 18 drugs in phase-III development, double what Merck had before. It also opens new markets for Merck: Schering-Plough earns 70% of its annual revenue outside the United States, including more than $2 billion from emerging markets.

The deal is the latest in a series of drug-industry marriages: Pfizer paid $68 billion in January to acquire Wyeth, and Roche recently boosted its offer to buy out Genentech to $46 billion.

Indonesia to sell carbon credits to conserve forests

Indonesia has applied to join a World Bank programme intended to help developing nations fight deforestation by selling tradeable carbon credits.

International climate negotiators are working to allow developing nations the right to sell some carbon credits if they clamp down on deforestation, which is responsible for roughly 20% of global greenhouse-gas emissions. The World Bank's US$300-million Forest Carbon Partnership Facility is designed to lay the groundwork for such an international agreement.

The programme already includes 25 countries, but Indonesia, the world's third-largest greenhouse-gas emitter, had remained on the sidelines until it applied in Feburary. The bank estimates that the country could earn between US$400 million and $2 billion selling credits for protecting forests.

H. BAKTIANTORO/REUTERS



World Bank incentives protect Indonesian forests.

Hawaiian waters yield corals with thirst for originality

Marine biologists have found seven new coral species, representing up to six new genera, in the deep waters of the Papahānaumokuākea Marine National Monument northwest of Hawaii.

The colourful bamboo corals, some of which were 4,000 years old, were discovered in November 2007 during a three-week exploration of the conservation area, which was created under a proclamation by former US president George W. Bush in 2006.



HAWAII DEEP-SEA CORAL EXPEDITION 2007/NOAA

The picture of this coral, from a seamount in water 1,750 metres deep, was taken by a submersible. The coral still lacks an official name. But members of biologist Les Watling's lab at the University of Hawaii at Manoa are calling it "Slurpisis". "It's built like a drinking straw, and 'isis' is the root name for the bamboo corals," explains Watling.

Drug maker found liable despite FDA warning label

Drug makers can be held liable for harm caused by their medicines, even if the drugs already carry federally approved warning labels, the US Supreme Court ruled on 4 March. The high court majority said that Diana Levine was entitled to a US$6.7-million award a state jury gave her, after she sued drug maker Wyeth of New Jersey.

A musician from Vermont, Levine had a Wyeth anti-nausea drug mistakenly injected into an artery, even though a label, approved by the US Food and Drug Administration (FDA), warned against such practice. Gangrene set in, and Levine's forearm was amputated.

The ruling means that FDA labels do not protect companies from liability in state courts, where they are often sued. It could affect many businesses that have sought increased federal regulation as a shield for liability.

For a longer version of this story, see: http://tinyurl.com/ckaubv.

NIH-funded primate centre accused of mistreatment

A Louisiana primate research centre came under federal investigation last week after an undercover probe by an animal-rights group alleged mistreatment of chimpanzees and monkeys.

The inquiry by the Humane Society of the United States also revealed that the National Institutes of Health (NIH) contracted the facility, the New Iberia Research Center, part of the University of Louisiana at Lafayette, to breed infant chimps for research.

Although the NIH has had a moratorium since 1995 on breeding chimps for research, an NIH spokesman says this doesn't cover a US$6-million contract that its National Institute of Allergy and Infectious Diseases (NIAID) made in 2002 with the centre to provide 4–12 infant chimps.

The US Department of Agriculture, which regulates primate facilities, announced an investigation on 5 March. The same day, legislators introduced a bill, the Great Ape Protection Act, seeking to ban all invasive research and testing on the estimated 1,000 research chimpanzees in the country.

Officials at the Centers for Disease Control and Prevention in Atlanta acknowledged that they received eight of the chimps for hepatitis C research under the NIAID contract.

Mars orbiter plans to map out methane plumes

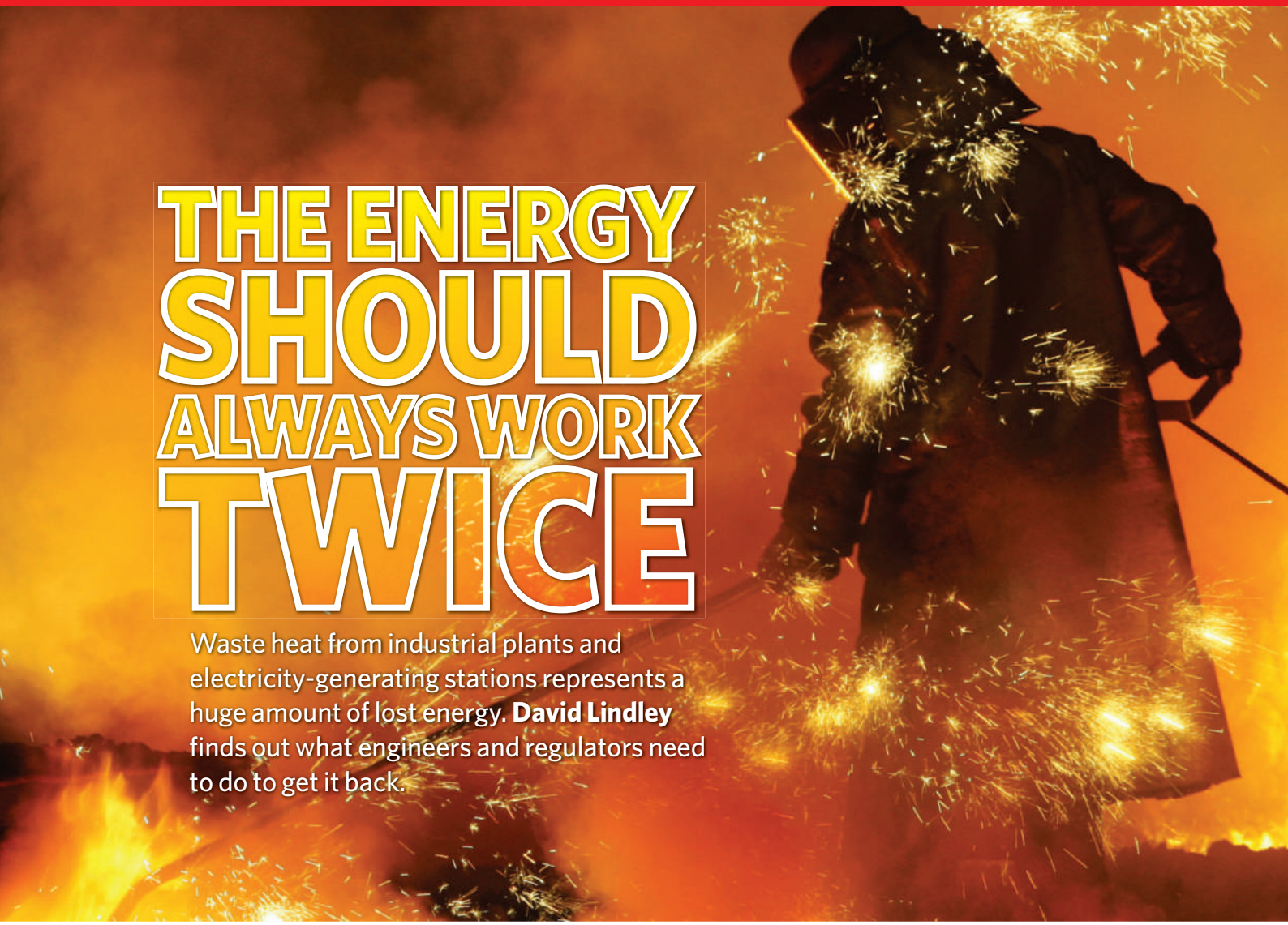
NASA officials said on 3 March that the agency may pursue a 2016 launch of an orbital mission to scout for methane hotspots on Mars, in preparation for a 2018 lander mission that would perform astrobiological tasks.

The orbital mission, called the Mars Science Orbiter (MSO), would produce detailed maps of atmospheric methane plumes recently hinted at by Earth-based observations and by older evidence from the Mars Express orbiter, which is funded by the European Space Agency (ESA).

In an effort to pursue joint Mars missions, ESA and NASA are looking at ways to get the MSO and ExoMars, a €1.2-billion (US$1.5-billion) rover also slated for a 2016 launch, on the same rocket.

© 2009 Macmillan Publishers Limited. All rights reserved

NATURE|Vol 458|12 March 2009



THE ENERGY SHOULD ALWAYS WORK TWICE

Waste heat from industrial plants and electricity-generating stations represents a huge amount of lost energy. **David Lindley** finds out what engineers and regulators need to do to get it back.

V. KESSLER/REUTERS/CORBIS

The power flows freely at West Virginia Alloys. This factory, tucked into the Kanawha River valley about 50 kilometres upstream from Charleston, West Virginia, is the second largest consumer of electricity in the state. It routinely pulls more than 120 megawatts of power from the grid, directing it into five electric arc furnaces that heat quartz to about 1,500 °C to make high-grade silicon for computer chips, solar cells and other uses.

Then the plant takes the furnaces' 800 °C exhaust, which still contains much of the original energy, and sends it billowing out of the smokestacks.

Unfortunately, West Virginia Alloys is not unusual, says Thomas Casten, an industrialist who has been preaching against this kind of energy waste for 30 years. Consumers around the world have embraced personal energy conservation in compact fluorescent light bulbs, home insulation and hybrid vehicles. But the industrial sector — including much of the electric power industry itself — continues to waste energy at a staggering rate. Thanks to generations of cheap energy, industrial process designers have typically had neither the expertise nor the economic incentive to do otherwise.

Now, however, West Virginia Alloys is one of a small but increasing number of companies trying to change their throwaway habits. By 2011, equipment being installed by Recycled Energy Development (RED) in Westmont, Illinois, a company that Casten co-founded in 2006, will start converting the plant's waste heat into electricity at a rate of more than 40 megawatts. That's enough to reduce the plant's power consumption by a third.

Such opportunities are everywhere, says Casten. There are 300 or so silicon plants in the world; each of them could repeat the West Virginia Alloys' story. Then there is steel. About half the steel in the United States is made by recycling scrap in arc furnaces; each furnace could recycle perhaps 15 megawatts of electricity. And for the manufacture of tyres, huge amounts of carbon black are used as a toughening agent for rubber. Carbon black is made by the partial combustion of tars, and the heat wasted could generate about 1 gigawatt of electricity.

Improving the energy efficiency of traditional fuel use is at least as good a way of protecting the environment as pushing for greater use of renewables, says Ahmed Ghoniem, a mechanical engineer at the Massachusetts Institute for Technology's Center for 21st Century Energy in Cambridge. Burning fossil fuels accounts for around 85% of the world's energy use, he says, so that a 20% increase in efficiency, "which is not unachievable", would reduce pollution and emission of greenhouse gases as much as would doubling the global proportion of 'green' energy generation — a goal that's years if not decades away, he adds. And in the United States alone, Casten says, energy recycling could offset the equivalent of about 200 gigawatts of electricity generation — about one-fifth of the nation's total capacity.

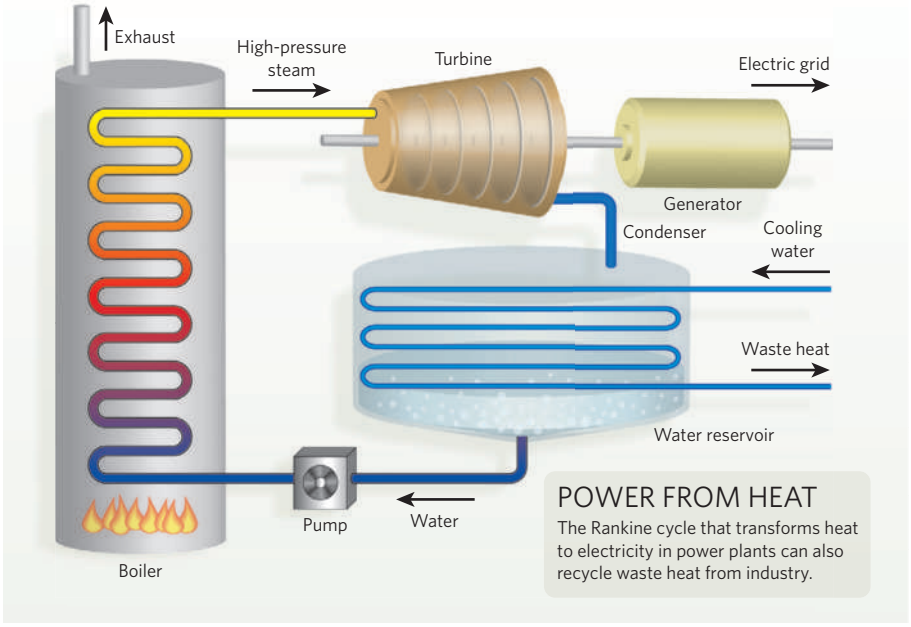
The challenge will be to get there — through a labyrinth of technological issues, costs, national energy policies and the often byzantine regulatory structures that govern electric power generation. "It's inordinately complicated to put all this together," says Casten.

Technology old and new

The good news for Casten's company and others in the energy-recycling field is that more and more potential customers are starting to pay attention. After last summer's spike in the price of oil, there is a widespread conviction that the price of energy has nowhere to go but up, however low oil prices might have fallen at the moment.

© 2009 Macmillan Publishers Limited. All rights reserved





POWER FROM HEAT

The Rankine cycle that transforms heat to electricity in power plants can also recycle waste heat from industry.

The bad news is cost: the price tag for retrofitting waste-heat recovery technology into a plant that was not designed for it tends to send potential customers into shock. The West Virginia Alloys installation will cost about US$100 million. That is why Casten's RED, for one, supplies both the capital and the expertise for installing energy-saving systems, in return for an agreement with the company to share the benefits. Plant managers don't have the knowledge to tackle the task, says Casten. "If they look at doing this on their own they will tend to overestimate the risks and underestimate the rewards."

Still, the technology itself is straightforward; in many cases, waste heat can be turned into electricity in exactly the same way that most power plants do it. A standard plant starts by taking heat from some primary source — coal, gas, biomass, enriched uranium or even concentrated sunshine — and using it to vaporize water into high-pressure steam. The steam is then directed through a turbine: essentially an arrangement of high-tech fan blades that are forced to spin by the steam's forwards motion. The turbine's shaft rotates an electrical generator, which produces the electricity. And the steam coming out on the far side is passed through a chiller that condenses it back into water again.

This Rankine cycle — named after William Rankine, the nineteenth-century Scottish engineer who devised it — is a practical and efficient way to turn a temperature difference into useful power (see 'Power from heat'). But as the French engineer Sadi Carnot showed in 1824, there is a fundamental limit on how efficient any such cycle can be. The bigger the temperature difference, the bigger that maximum efficiency is; thus engineers' frequent references to 'high-quality' heat, by which they mean that the source is much hotter than its surroundings. But no matter how good the technology, there will always be some heat energy left over. In the case of the Rankine cycle, this is primarily the energy given up by the steam as it condenses back into water. In principle, further mechanical work can be extracted from leftover heat, as long as there is at least some temperature difference with the surroundings. But usually, says Ghoniem, "because it's low quality you have to work a bit harder to get that energy out".

That's why energy-recycling companies go after high-quality heat first. At West Virginia Alloys, for example, the 800 °C exhaust heat from the furnaces will be used to boil water, which will then generate electricity via a standard Rankine-cycle system. And there are planty of other examples, says Casten. He vividly remembers visiting a lime plant in Nevada that has enormous cylindrical kilns fed by limestone, pulverized coal and oxygen. "You can feel the heat 30 feet away," he says. Large-scale industries such as ore-smelting and metal-refining have conventionally pumped out energy that does nothing except heat the surrounding air.

"Separate generation of electricity and heat is utter madness."
— Thomas Casten



RED

Tempting as these targets may be, there are many more opportunities to reclaim lower-quality waste heat given off in chemical manufacturing, paper making and countless other industries. RED is looking at the cavernous ovens used to dry great sheets of plasterboard (drywall) — an energy consumption that represents 80% of the plasterboard's cost. The exhaust temperature is less than 200 °C, making it hard to run a water-based Rankine cycle with any efficiency. But it is now possible to buy off-the-shelf Rankine-cycle systems that replace water with more volatile liquids such as freon, propane or butane, which can extract mechanical work from more modest temperature differences.

At the high-tech frontier, meanwhile, are waste-heat recovery systems that rely on solid-state thermoelectric materials that directly generate electricity from a temperature gradient. Commercially available thermoelectric devices can attain 15–20% of the Carnot efficiency for a given temperature difference, says Jeff Snyder, a materials scientist at the California Institute of Technology in Pasadena, making them both less efficient and more expensive than a Rankine device doing the same job. Nonetheless, their compact size and lack of moving parts make thermoelectric generators an attractive option for the automotive industry. BMW and Volkswagen, for example, have both announced prototype systems that would create electricity from the engine's exhaust heat.

Researchers are exploring more-exotic chemical compositions and materials engineered to have nanostructures to decrease thermal conductivity, slow the flow of heat and make its energy easier to capture. Any heat that

© 2009 Macmillan Publishers Limited. All rights reserved





T. FALBE-HANSEN

BIOPRESS

About half of Denmark's electricity comes from combined heat and power plants (such as these in Viborg, left, and Amager) that provide both electricity and heat to houses and offices.

flows all the way through the material is lost, says Snyder. Such systems probably won't be able to compete with Rankine-cycle technology in a factory setting, he says, but it's realistic to think of adding thermoelectric generators to vehicles at a cost a few hundred dollars.

The biggest offender

Globally, one of the most glaring inefficiencies comes from the burning of fossil fuels to generate electricity. About two-thirds of the world's electricity is made this way, according to a 2005 tabulation from the US Energy Information Agency, and the conversion efficiency in typical coal- or oil-fired plants is about one-third.

Newer electric power plants running on natural gas or on 'syngas', a mixture of mostly hydrogen and carbon monoxide created by gasification of coal, can achieve greater efficiencies through a combined cycle system, in which combustion of the gas directly drives a first set of turbines, while exhaust heat from that stage generates steam that drives a second set of turbines. Combined cycles deliver efficiencies of 55% or more, says Ghoniem. Gasification is a cleaner but more expensive use of coal, whereas natural gas is mostly used in 'peaker' plants that can be turned on and off quickly as changing demand requires. Combined cycle systems have therefore yet to make a significant improvement in the average efficiency of converting fossil fuels to electricity.

Still, no method for generating electricity by burning a fuel can avoid substantial inefficiency through the loss of heat — which puts utilities in the ludicrous position of venting huge amounts of waste heat to produce electricity that is often turned back into heat somewhere else, via arc furnaces, hair driers and home heating systems. So why not use that waste heat directly?

"Separate generation of electricity and heat is utter madness," says Casten, who likes to point out that the plant built by Thomas Edison in Manhattan in 1882 to supply electricity to the burgeoning city also sent its waste heat through pipes to warm nearby buildings. Such a combined heat and power (CHP) system remains in place in New York City, operated by Consolidated Edison and feeding steam heat to about 100,000 buildings.

The emphasis is on 'nearby', however; heat doesn't travel as well as electricity does. So in the United States, where the power system evolved through the twentieth century towards large power plants located in remote areas, this municipal deployment of CHP remains rare. But other countries do things differently.

Perhaps the most notable example is Denmark, which has pursued national energy policies that have put it at the forefront of energy efficiency as well as use of renewable sources, principally wind power. About half of Denmark's electricity comes from CHP plants — a few large ones, which switched from oil to coal following the oil crisis of the 1970s, and thousands of small to medium units, serving smaller communities and mostly running on natural gas. Some local units burn wood chips, straw, other biomass and household waste. "Everything that can be burned is burned," says Per Lund, head of system development at Energinet.dk, the state-owned corporation in Fredericia that runs Denmark's electricity grid.

"The regulatory system allowed private investment, and the network people couldn't stop them."
— Thomas Ackermann

Part of Denmark's success with CHP is a matter of geography: its large CHP plants sit close to the shore and use cold sea water for cooling. As larger temperature differences allow greater efficiency, these plants achieve efficiencies of more than 90% in terms of the electricity and useful heat obtained from their fuel. Over the past 25 years, Denmark's energy consumption has remained essentially constant even as the country's gross domestic product (GDP) has grown by 75%, giving it the European Union's lowest 'energy intensity', measured in terms of energy consumption per unit of GDP.

But Denmark's success also derives from a mix of tight regulations, government investments and open-market responses. For example, says Lund, there was significant national spending to install the buried pipes that distribute hot water around communities of as few as 500 houses. And in CHP districts, which include about 60% of the country's population, residents must draw their domestic heat from the centralized system. But at the same time, private investors who see a profitable opportunity in providing additional energy are free to propose building a new CHP unit. Likewise, industries that invest in equipment to capture and divert their waste heat can recoup some of their costs by selling it to the local CHP district.

Keeping up with the competitors

In the United States, Kathey Ferland, project manager at Texas Industries of the Future, an advisory organization in Austin, points out that energy-intensive oil refineries and chemical plants in Texas have long made use of CHP to achieve internal efficiencies. "What mostly seems to motivate them is making sure that they are at the least doing what their competitors are doing," she says.

However, in Texas, Ferland says, "a lot of the really good economic sites are done", and for smaller operations the economics of CHP can be hard to work out, or outright unfavourable. For example, a plant that uses CHP to meet its basic power needs may greatly reduce its consumption of off-peak electricity — a significant fraction of which comes from wind power in Texas — while still drawing on the grid when its needs are higher. CHP, therefore, competes disproportionately against the cheapest and

© 2009 Macmillan Publishers Limited. All rights reserved

most environmentally friendly electricity, without eliminating the need for conventional power plants.

So an electricity-pricing system with higher daytime prices to discourage peak consumption can create a "misincentive" by acting against other desirable goals, says Chris Marnay of the Lawrence Berkeley National Laboratory in California. "The form of the technologies that are successful is determined at least as much by the regulatory situation as it is by the engineering reality." In one study of medium-scale commercial buildings, Marnay and his colleagues found that CHP made economic sense for a hypothetical building in California, but not for the same building in New York, mainly because of differences in the electrical tariff structure.

National oversight

These regional variations arise because electricity regulation has traditionally been in the hands of the states. The 1978 Public Utility Regulatory Policies Act modified that picture, obliging US electric utility companies — at that time, a collection of local and regional monopolies — to buy electricity from independent producers. Oversight was put in the hands of individual states, which took on the task with varying degrees of enthusiasm, and the terms of the act gave utilities considerable influence over the price they were supposed to pay for independently produced electricity.

The Danish government, by contrast, mandated expansion of CHP as a national policy beginning in the 1980s, requiring grid operators to pay for the power generated at a set rate. The guaranteed ability to sell electricity stimulated the construction, by private entities, of relatively small-scale CHP plants. The same strategy lies behind the rapid growth of wind and solar power in Denmark and also in Germany. In both cases, grids were legally obliged to buy renewable power. Wind power, which now accounts for about 20% of Denmark's electricity supply, began as a grass-roots movement of "happy amateurs", says Lund, with farmers installing turbines made by the same companies that forged ploughs and built water tanks. Today, however, wind turbines are huge business, manufactured by a handful of large companies.

At the time, power engineers accustomed to working with grids that transmitted power over wide areas from a handful of enormous generating plants doubted that the new system could work reliably, says Thomas Ackermann, chief executive of Energynautics in Langen, Germany, an energy-consultancy company. A number of Danish studies argued that distributed generation — with electricity coming from many smaller-capacity suppliers — couldn't be done. And "it wasn't done, in the sense that anyone actively planned it", he says. "Instead, the regulatory system allowed private investment, and the network people couldn't stop them."

As a result, Denmark is moving towards an increasingly distributed generation system, in which electricity doesn't simply flow in a 'top-down' manner, but can also move in a 'bottom-up' way from multiple small generating units. This required many changes to the way power moves through transformers from low-voltage to high-voltage parts of the network. A unified national policy of private investment aided by government subsidy spurred the construction of moderate-capacity electricity plants, and the grid operators were able to pass on costs to customers through tariffs. But in many countries, such transformations falter on the question of who will pay for them, says Ackermann.

That is typically the case in the United States, where efforts to deregulate the old monopolies have happened piecemeal. "One of the first problems is that people don't understand each other, everyone has a different mindset, everyone has a different way of looking at the system," says Ackermann. A common question is whether a distributed generating system with many small seasonally variable suppliers can ensure that peak demand will be met.

> **"The US system looks like our system did maybe 25 years ago."**
> — Per Lund

But, says Ackermann, in a reversal of the standard cliché about the free-market United States and socialist Europe, "we don't talk about guaranteeing power from wind or other sources all the time — we just believe that the market will cope". In Germany, he says, peak demand is about 74 gigawatts and the market can supply about 120 gigawatts, so that as demand waxes and wanes, more expensive producers enter and leave the market, and the price of electricity rises and falls. "The system is one of the best in the world in terms of reliability," says Ackermann. He argues that the greater regulatory burden in the United States has inhibited development of diverse generating capacity and of the grid, so that the infrastructure that gets built is the bare minimum. "The US system looks like our system did maybe 25 years ago," says Lund.

Change we need?

But change is in the air. In his Senate confirmation hearing, Steven Chu, former director of the Lawrence Berkeley National Laboratory and now US energy secretary, declared that "more efficient use of energy in the United States is the one big factor that can most help us reduce our dependency on foreign oil". The economic stimulus package signed by President Obama on 17 February includes a modest federal subsidy for CHP and recycled-energy projects. "This is the first federal inducement of generation efficiency since 1986," says Casten, although he adds that the bill promises much greater support for renewable energy sources, "which is not as good a way to spend money". He hopes that when Chu talks about efficiency, he doesn't just mean better appliances and more insulation. "The bottom line is that we will not have a prayer of mitigating climate change unless we address the efficiency of generating thermal and electric energy." ■

David Lindley is a freelance science writer based in Alexandria, Virginia.

See Editorial, page 125.



R. HAMILTON SMITH/CORBIS

Each furnace used to turn scrap metal into steel could be used to recycle about 15 megawatts of electricity.

© 2009 Macmillan Publishers Limited. All rights reserved

NATURE|Vol 458|12 March 2009



Of owls, larks and alarm clocks

Could out-of-sync body clocks be contributing to human disease? **Melissa Lee Phillips** reports.

Ten years ago, researchers reported on three families with an extreme 'lark' problem. Larks are people who naturally wake up early in the morning, and are the opposite of 'owls', who wake up and go to sleep late. But many of the members of these families had particularly acute larkness, waking up on average around 4 a.m. and falling asleep around 7.30 p.m.[1]. The researchers later found that the families, who were diagnosed with familial advanced sleep-phase syndrome (FASPS), all carried mutations in a gene called *PER2,* which is involved in setting the human body clock[2].

By some estimates, more than half of the population in industrialized societies may have circadian rhythms that are out of phase with the daily schedule they keep. Such people are said to have 'social jet lag' — a term coined by Till Roenneberg of Ludwig Maximilians University in Munich, Germany. Some of these are larks, some owls, and some have pretty standard human rhythms that are disrupted by shift work or travel. In modern societies, circadian-rhythm disruptions can arise from simply spending too much time indoors, something that can make such workers decidedly "owlish", Roenneberg says. Even the one-hour time change made by many countries at this time of year can take some adjusting to.

If larks and owls are forced to follow normal schedules, they run into all kinds of problems with disabling insomnia and sleepiness. But disrupted rhythms could have graver consequences than that. In 2007, an expert working group at the World Health Organization's International Agency for Research on Cancer in Lyon, France, concluded that "shift-work that involves circadian disruption is probably carcinogenic to humans"[3], after reviewing the existing evidence. Equally strong conclusions have been drawn from evidence that links circadian-rhythm problems to psychiatric disorders, metabolic syndrome and a range of other illnesses.

> **"You have to view your body as a whole collection of different clocks."**
> — Joseph Takahashi

Researchers now are working to understand those links. Some suspect that health problems arise from a third kind of jet lag — one that arises when the circadian rhythms in different body tissues lose synchrony with each other. In 2006, the European Commission started funding EUCLOCK, a €16-million (US$20-million), five-year project involving some 34 researchers whose goal is to understand how the circadian clock synchronizes with cycles in the environment. In particular, the researchers are trying to work out what type of schedules are the healthiest fit for individuals' biological clocks. "People have been researching this for 50 years," says Anna Wirz-Justice at the Centre for Chronobiology in Basel, Switzerland, "but I think the methods are only now coming up to address this properly."

In nearly all organisms, patterns of biochemistry, physiology and behaviour oscillate with the daily cycles of light and dark, often with near-perfect timing. People forced to live in a 28-hour cycle still show fluctuations of almost exactly 24 hours in their core body temperature and levels of the hormones melatonin and cortisol[4]. In mammals, many of these cycles are directed by a 'master clock' in the brain's hypothalamus called the suprachiasmatic nucleus (SCN). The SCN receives information from the retina about light and coordinates rhythmic cycling of gene expression in the rest of the brain and body through neural signalling and hormones. Cycles of gene expression in

ILLUSTRATIONS BY MATT MURPHY

© 2009 Macmillan Publishers Limited. All rights reserved

the gastrointestinal tract, for example, ensure that digestive acids and enzymes are produced at the appropriate times.

In the past decade, researchers have identified some of the genes responsible for timekeeping, showing that expression of *CLOCK*, *BMAL1*, *PER* and *CRY* rises and falls over 24 hours in the SCN and elsewhere, and that mice with mutated versions of these genes abandon the 24-hour cycle for one that is shorter, longer or has no pattern at all. Such animals were "the key development that brought the field to its present exciting position", Wirz-Justice says, because it suggested that these were 'master genes' directing the clock and the physiological processes that follow it.

Typecasts

But for some individuals, the cycles of gene expression and behaviour do not adhere well to the cycles of night and day. The tendency to be a lark, an owl, or somewhere in between is referred to as individual's 'chronotype', and although it may shift over the course of a person's lifetime — adolescents and young adults tend to be more owl-like than either children or older adults — it doesn't usually change in comparison with peers and it is thought to be determined largely by genes. FASPS was the first human circadian disorder linked to a mutation in a specific clock gene. Not all such genes have been easy to find: researchers have found no simple mutation that accounts for people with the owlish delayed sleep-phase syndrome (DSPS), who can have trouble falling asleep before 6 a.m. and waking up before 2 p.m..



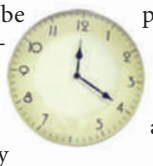
People with DSPS and FASPS often also have depression[5,6], and this and other psychiatric conditions, such as bipolar disorder and schizophrenia, are commonly associated with abnormalities in circadian rhythms. "The vast majority of people with major depression have sleep abnormalities and interestingly it can be that they sleep too much or they have insomnia and can't sleep," says Colleen McClung of the University of Texas Southwestern Medical Center in Dallas. This connection raises a cause and effect question: are the circadian-rhythm disorders causing the depression or the other way around? In 2007, McClung and her colleagues found some support for the former idea when they studied mice that lack a working *Clock* gene and, they observed, exhibit symptoms of mania and hyperactivity that can be reversed by the mood stabilizer lithium[7]. "Every way we test them, they look like bipolar patients in the manic state," McClung says.

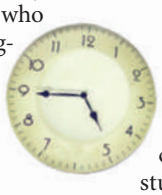
Circadian rhythms are known to affect the most basic of metabolic pathways, including protein synthesis, glycolysis and fatty-acid metabolism. And many patients who have circadian-rhythm disorders caused by shift work also develop gastrointestinal and metabolic disorders such as glucose intolerance, diabetes and high blood pressure, says Theodore Bushnell, a neurologist at the University of Washington Medicine Sleep Disorders Center in Seattle. "It seems like there's nobody who just has a shift-work issue."



The conclusion that shift work can be carcinogenic has grown largely from epidemiological work. A nationwide study in Denmark, for example, found that women who work mainly at night for at least six months are 1.5 times more likely to develop breast cancer than those who work regular hours[8]. Researchers suggest that the raised cancer risk could be because these people's cells start to divide at the wrong time and run amok, an idea supported by some cell-culture studies.



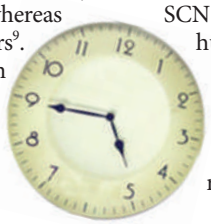




So far, these links haven't provided the details that biologists would like. Much like diet or stress, circadian rhythms affect so many cellular and physiological functions that it is extremely difficult to pinpoint a mechanism by which a given alteration to these rhythms could contribute to a particular disease. Much of the current work involves genetic association studies, in which scientists look for gene variants that pop up more often in people affected with a disorder than in those without it. Some such studies for psychiatric diseases, for example, have pulled out variants in clock genes. But researchers have also been making progress by dissecting the machinery of the various human body clocks.



Since the late 1990s, it has become apparent that the body has 'peripheral circadian oscillators' in tissues outside the SCN. These peripheral clocks receive input from the SCN, but are also influenced by other timekeeping signals. In 2004, a team led by Joseph Takahashi of the Center for Sleep and Circadian Biology at Northwestern University in Evanston, Illinois, measured cycles of gene expression in different mouse tissues in culture when placed in constant darkness. He found that the kidney cells followed a clock of 24.5 hours, whereas the corneal cells ran at about 21.5 hours[9]. "We used to think the clock was just in the brain — it was a neural process and the body just passively followed," Takahashi says. "But that's clearly not the case. You have to view your body as a whole collection of different clocks."



Like Greenwich Mean Time for the body, the SCN serves as a reference point for the peripheral clocks, and they usually run in sync. But sometimes — after stepping off a long-distance flight, for example — the synchrony breaks down. When simulating jet lag in rats, researchers at the University of Tokyo found that the SCN resets to local time in around one day based on light signals, but peripheral oscillators can take more than a week to adjust[10]. "So, during jet lag, your body is literally completely out of sync," Takahashi says. "Each organ shifts at a different rate."

Could it be that loss of synchrony between the body's tissues underlies some of the health problems seen in circadian-rhythm disorders and shift workers? To test this, researchers have used genetic tools to create an artificial mismatch between the central clock and a peripheral one. In one study[11], researchers knocked out the gene *Bmal1* in the liver of mice, effectively disabling the clock in that organ. When the animals developed low blood sugar for parts of the day, the researchers knew that the brain clock was not sufficient to maintain glucose levels — a working peripheral clock was needed too.

The SCN is synchronized during the daytime by light, whereas liver metabolism is also synchronized by food intake, Roenneberg says. So perhaps shift-workers have gastrointestinal problems because their liver and intestinal tracts are gearing up for a meal at the wrong time. "You can easily imagine that this is not exactly optimal for the system," he says.

Out of time

A mismatch between central and peripheral clocks has been linked to other health conditions too. Ongoing work by Christian Cajochen at the Centre for Chronobiology in Basel, suggests that the peripheral clocks of women with depression are not as well linked to sleep–wake cycles as they are in those without the condition. "Women with depression have a greater degree of variability in the timing of different physiological and endocrine rhythms," Wirz-Justice says.

Some researchers in the EUCLOCK consortium are planning to work out how the SCN is able to keep peripheral clocks and human physiology in time. In some experiments, mice will be put onto 'shift-work' schedules in which they are forced to 'work' and eat during the day rather than at night. The researchers will then examine how the

© 2009 Macmillan Publishers Limited. All rights reserved



schedule affects behaviour and the synchronization of various local body clocks.

In fact, peripheral oscillators have already proved to be a useful research tool. Last February, researchers from Steven Brown's lab at the University of Zurich and Achim Kramer's group at Charité Hospital in Berlin identified 11 larks and 17 owls based on a 'morningness–eveningness' questionnaire and then measured their molecular rhythms from the expression of *Bmal1* in their skin cells[12]. The team wanted to find out why larks and owls naturally adopt the schedules that they do. For some people, the researchers found what they expected: cells from larks showed shorter periods than those from owls. But they also found that about half of the larks and owls actually had 'normal' circadian period lengths.

> **"The vast majority of people with major depression have sleep abnormalities."**
> — Colleen McClung

The researchers found that in this group, the owls had skin clocks that were more difficult to reset than those of people with more typical schedules, and that the larks had clocks that were easier to reset. This suggests that individual differences in chronotype result not just from innate differences in circadian period length but also from differences in how easily a person's rhythms can be synchronized to the night–day cycle — and that some larks and owls have clocks that are not reset normally each day. So perhaps these people's peripheral clocks stray from the central one easily.

Therapy options

With few firm mechanisms to go on, the question now is how to go about 'treating' circadian-rhythm disruptions. Intense-light therapy has been used to shift undesirable sleep schedules back to a more normal pattern. And Wirz-Justice says that researchers are working on other behavioural or pharmaceutical ways to alter circadian rhythms. Ramelteon, for instance, is a drug used to reduce the effects of insomnia by mimicking the action of melatonin — a hormone that tells the body it is sleep time. But no one knows yet whether such interventions will also prevent or reduce some of the other health risks associated with circadian-rhythm disruptions. Roenneberg is not convinced that it would be easy to solve even conventional jet lag with drugs. "The trouble is that we are still much too naive to have pharmacological success without messing something else up in the system," he says.

That means it could be up to individuals to consider their chronotype before choosing their schedule, or making other decisions that might affect their health. EUCLOCK researchers are developing ways to measure clock-gene expression from cheek swabs or other tissue that could provide a quick laboratory read-out of a person's chronotype. (Roenneberg says that so far these results agree with self-reported chronotype based on questionnaires.) Chronotype might be taken into account when administering drugs or medical tests: performing a blood test at 8 a.m. can yield completely different results on a morning person, who has been up for hours, than it will on an evening one. "You can get very different clinical blood values based on their chronotypes without anything being meaningfully different, because in one person, the system is up, and in the other, it's still down," says Roenneberg.

Working with EUCLOCK and another consortium called CLOCKWORK, Roenneberg and his colleagues are developing a computer model that works out a person's ideal schedule on the basis of his or her chronotype and sleep needs. The group is starting to test the model's predictions by asking people of each chronotype to try out a certain schedule in their daily lives, and measuring its effects in a range of physiological, behavioural and cognitive tests. They will then use their findings to refine the model. By repeating this process several times, Roenneberg says, "we are confident that we will finally put some sense into how to do shift-work properly — that is, with the fewest side effects on health".

Being a lark or owl should not, in itself, be a problem, Roenneberg adds. "Chronotype per se should have no health effects whatsoever. Health effects come from having to live against one's own clock." ■

Melissa Lee Phillips is a freelance writer based in Seattle, Washington.

1. Jones, C. R. *et al. Nature Med.* **5,** 1062–1065 (1999).
2. Toh, K. L. *et al. Science* **291,** 1040–1043 (2001).
3. International Agency for Research on Cancer *Lancet Oncol.* **8,** 1065–1066 (2007).
4. Czeisler, C. A. *et al. Science* **284,** 2177–2181 (1999).
5. Shirayama, M. *et al. Sleep Med.* **4,** 427–433 (2003).
6. Xu, Y. *et al. Nature* **434,** 640–644 (2005).
7. Roybal, K. *et al. Proc. Natl Acad. Sci. USA* **104,** 6406–6411 (2007).
8. Hansen, J. *Epidemiology* **12,** 74–77 (2001).
9. Yoo, S.-H. *et al. Proc. Natl Acad. Sci. USA* **101,** 5339–5346 (2004).
10. Yamazaki, S. *et al. Science* **288,** 682–685 (2000).
11. Lamia, K. A., Storch, K. F. & Weitz, C. J. *Proc. Natl Acad. Sci. USA* **105,** 15172–15177 (2008).
12. Brown, S. A. *et al. Proc. Natl Acad. Sci. USA* **105,** 1602–1607 (2008).

© 2009 Macmillan Publishers Limited. All rights reserved

CORRESPONDENCE

These letters respond to the two Commentaries discussing 'Should scientists study race and IQ?' — 'NO: Science and society do not benefit' by Steven Rose (*Nature* **457,** 786–788; 2009) and 'YES: The scientific truth must be pursued' by Stephen Ceci and Wendy M. Williams (*Nature* **457,** 788–789; 2009).

The belief that genes cannot be changed is now outdated

SIR — The relevance of race-and-intelligence research is obvious. The most troubling feature of the world economy today is not the financial crisis, but the enormous difference in wealth, technological and cultural creativity, technical know-how and social organization between countries. The per-capita gross domestic product (GDP) is 30 times higher in the United States than in most African countries.

We do not know the reasons for the great divergence that created these inequalities, and without this knowledge we shall never be able to reverse them. Genetics is one hypothesis. It is also the one that can be tested most easily, with the aid of genome-wide association studies and whole-genome sequencing.

People have a strong desire to believe in a just world. The increasingly atheist intellectuals of the twentieth century gave up their belief in divine justice only to substitute for it the more 'scientific' belief that Mother Nature is just: that nobody is disadvantaged by his or her genes, and the undeniable inequities in the world are exclusively man-made. But the belief that individual differences in intelligence are unaffected by genes was shot down by a barrage of behaviour–genetic studies during the 1980s. Now science is threatening the very last bastion of the just-world belief: that Mother Nature cannot be blamed for inequities between entire nations and population groups.

By not investigating the race–intelligence link, we not only perpetuate ignorance and the prejudice that thrives on ignorance. We also deprive ourselves of the possibility to tackle the existing inequalities, first by a judicious development policy and — should genetic differences indeed be important — by eventually changing the allele frequencies of the offending genes. We should not get stuck in the twentieth-century assumption that environments are changeable but genes are not. This will no longer be the case in the twenty-first century.



Investigations into the genetic aspects of race and intelligence are part of a wider enterprise in basic science: the study of the recent and ongoing evolution of human intelligence. This whole area of basic research will have to be scrapped if we refuse to study allele frequencies of cognition-related genes in human populations.

Gerhard Meisenberg Department of Biochemistry, Ross University School of Medicine, Picard Estate, Dominica, Windward Islands
e-mail: GMeisenberg@rossmed.edu.dm

Identifying adaptive differences could provide insight

SIR — I agree with others who have said we should not expect different subgroups of the human population, evolving independently, to arrive at exactly the same place in terms of cognitive abilities. This makes no more sense than expecting different populations to end up identical in skin colour, stature, metabolism or other aspects that are easily understood as adaptations to different environments.

So, given that we have logical reason to hypothesize about differences in cognitive abilities, why would we expect to measure these by using a single number such as IQ, which suggests there must be a hierarchy of cognitive function? The prediction surely is that each population will adapt to be better at the particular cognitive tasks that are most important for survival in its own environment. If this is the case, then identifying these (potentially adaptive) differences in cognitive ability, and searching for associations with genetic variants, could provide fascinating insights into how our brains work.

However, this is worlds away from measuring IQ of different 'race' groups in order to make claims about genetics and intelligence. There may be some value in these rather simplistic studies of race and IQ. But they do nothing to answer the scientific question of the genetic basis of intelligence and can easily be hijacked by individuals to advance their own prejudices.

Kathryn Holt Wellcome Trust Sanger Institute, Hinxton, Cambridge CB10 1SA, UK
e-mail: kh2@sanger.ac.uk

The arrogance of trying to sum up abilities in a number

SIR — Rose, Ceci and Williams ask whether scientists should study race and IQ. The problems with such studies seem to result, as they have done since the 1930s, from the near impossibility of defining the word 'intelligence'.

The introduction of IQ tests has always seemed to be one of the best examples of the great political and social harm that can be done by the mind-boggling arrogance of scientists who think that they can sum up human abilities in a single number.

The same sort of problem exists with the present generation of bibliometrists. So far they haven't managed to inflict as much damage but, left to themselves, they will.

David Colquhoun Department of Pharmacology, University College London, London WC1E 6BT, UK
e-mail: d.colquhoun@ucl.ac.uk

Is poverty better explained by history of colonialism?

SIR — It is difficult to see the Commentary by Ceci and Williams as anything other than a whitewash. Decisions about what kinds of scholarly research questions and methods are considered worthy of attention and funding are fundamental to modern science.

Stupid science and evil science — both of which exist — should not be permitted to coexist casually alongside the normative intellectual activities we admire. It is the role of scientists, as gatekeepers, to distinguish among them, to identify them for non-specialist audiences and to repudiate the intellectually impoverished elements. Any science that fails to do this, that takes all work to be of equal stature, necessarily calls into question its own standing as a scholarly enterprise.

The study of an organic basis of intelligence is not itself threatening. But it does not explain economic stratification, poverty and illiteracy rates any better than the history of slavery and colonialism does. Anyone who thinks that it can is obliged to confront and acknowledge the political nature of the science

© 2009 Macmillan Publishers Limited. All rights reserved

they are engaged in, and must be prepared to defend it on that basis.

Racism is a political act, and scientific racism is simply the recruitment of the trappings of science in pursuit of its ignoble goals. If scientific racism has a place in science, it debases the entire enterprise.

Jonathan Marks Department of Anthropology, University of North Carolina at Charlotte, Charlotte, North Carolina 28223, USA
e-mail: jmarks@uncc.edu

Would you wish the research undone?

SIR — Rose offers a piece of moral advice in his Commentary. To instruct humankind that they should not do something is a serious undertaking: it should be clear what is enjoined and the consequences acknowledged.

Rose seems to argue that, by about 1975, it should have been clear to everyone that the question of genetic causes behind the black/white IQ gap was unanswerable. From that date, there was no excuse for Arthur Jensen or anybody else to persist.

I invite everyone to search the social-science literature of the past 34 years and ask whether or not they really wish that everything on the subject, pro or con, was missing. I thought that the first replies published to rebut Jensen were pathetic, and still do. Others must judge the value of my subsequent book *Race, IQ and Jensen* (Routledge & Kegan Paul; 1980); the theory of intelligence that limits *g* — the 'general intelligence factor' — to explaining individual differences; the Dickens–Flynn model of the interaction of genes and environment; the analysis of the achievements of Chinese-Americans, and so forth. More distinguished contributions have been made than those that I have offered.

Imagine that IQ tests showed black American children having, on average, a 15-point IQ advantage over white children. How many of us would denounce that finding as meaningless? Or would we be waving it like a banner, as evidence that only racists would attempt to dismiss? The scientifically respectable has a way of shifting, depending on how one thinks the debate is going.

As the philosopher John Stuart Mill points out, when you assert that a topic is not to be debated, you are foreclosing not some narrow statement of opinion on that topic, but the whole spiralling universe of discourse that it may inspire. Mill thought that only someone so self-deluded as to think his own judgement was infallible could wish to circumscribe an unpredictable future in this way.

Rose should be very certain he is correct. If not, and if he converts the rest of us, only Jensen and those of his persuasion will publish; and they will win the minds of students because the rest of us have all adopted a policy of unilateral disarmament.

Jim Flynn Box 56, University of Otago, New Zealand
e-mail: jim.flynn@otago.ac.nz

Measured intelligence is a product of social processes

SIR — 'Eugenics' is a word that most people believe has been consigned to history. The term makes us think of the horrors of Nazi Germany, or perhaps the sterilization of disabled people in places such as the United Kingdom, the United States and Sweden (see T. Shakespeare *et al. Genetic Politics: From Eugenics to Genome* New Clarion Press, 2002).

However, eugenic thinking also survives in contemporary education policy, in the belief that some children are simply 'brighter', 'smarter' or 'more able' than their contemporaries. In the US and UK systems, this assumption finds expression in schools' use of 'tracking' and 'setting by ability', which separates children (from as early as the age of five in Britain) and sets them on very different paths.



BETTMANN/CORBIS

Do race and gender affect performance in the classroom?

Yet there is no way of measuring the limits of capability. Every test ever invented (whether an IQ test or a driving test) assesses only how well a person is currently able to perform certain tasks. Measured 'intelligence' is a product of social processes, not a determinant of them. The results often become self-perpetuating. Black children and their white peers from poor backgrounds are consistently over-represented in the lowest-ranked groups, where they cover less of the curriculum, are taught by less experienced teachers and make slower progress.

It is time that we were liberated from the racist and regressive ideas that have become so intrinsically bound up in the notion of intelligence. By finally having the courage to admit that contemporary patterns of class, race and gender inequity are wholly a product of our own policies and priorities, we might finally begin the urgent business of dismantling such injustice, rather than seeking to excuse it as an act of nature.

David Gillborn Institute of Education University of London, 20 Bedford Way, London WC1H 0AL, UK
e-mail: d.gillborn@ioe.ac.uk

Don't fan the flames of a dead debate

SIR — Why study possible links between race, gender, genes and intelligence? In their Commentary, Ceci and Williams's response seems to be that it provokes good scientists to sharpen their critique of the poor science of those who ask that question, and that this advances the field. I agree; when, in the late eighteenth century, rival proponents of phlogiston and oxygen battled it out to account for combustion, the controversy was illuminating. But it was decisively settled, and within a decade phlogiston was dead. So too with race, gender and IQ. We had the debates some 30 years ago and, as Ceci and Williams agree, those who argued for genetic explanations for group differences (phlogiston theorists) lost. So why reignite a dead debate, unless it is to serve some sociopolitical, not scientific, end?

Ceci and Williams list some of those who have continued to insist that the issue is not closed, and defend them against the criticisms, and in some cases hostility, they have generated. To make their point, they invoke the spectre of Trofim Lysenko. The comparison is illuminating. Lysenko falsified experimental

© 2009 Macmillan Publishers Limited. All rights reserved

Food aid fails to provide a sustainable solution to hunger and poverty
Pedro A. Sanchez on cheaper ways to combat hunger, page 148

data in defence of politically expedient but mistaken theories, with damaging effects on Soviet agriculture. When in the 1940s his theories received Joseph Stalin's imprimatur, Soviet genetics, which had flourished through the 1920s and 1930s, was destroyed. But they have their comparison the wrong way round. In the present case it is the proponents of race–gender–IQ theories who are defending a mistaken but politically expedient theory, with potentially seriously damaging social consequences.

Ceci and Williams defend the principle of free speech; you should not be silenced for arguing for phlogiston, although your grant applications based on the theory are unlikely to be funded (except by Foundations for the Defence of Phlogiston, which certainly exist in the case of race, gender and IQ). But it is my impression that, far from being silenced, phlogiston theorists get quite a bit of time on air and in print. Freedom of speech should be defended, but it cannot be an unbounded freedom in cases where, more than just falsely crying "Fire!" in a crowded theatre, the speech provides the matches and petrol for those who are actively trying to set the place alight.

Steven Rose Open University, UK
e-mail: S.P.R.Rose@open.ac.uk

A useful way to glean social information

SIR — Rose concludes by noting, when it comes to research on race, gender and IQ, that "there is no valid knowledge to be found in this area at all". Is the worldwide increase in IQ not valid, and should its insights into the malleability of racial differences in intelligence be ignored? Is the superiority of women's mathematical grades unimportant? Does the convergence of black and white test scores not inform our understanding of how to achieve a more just society? All of these findings stemmed from research that Rose claims is invalid.

One-party science may appease some now, but will eventually stymie scientific understanding. The recent renaissance of research on Lamarckian modes of genetic transmission shows that not all widely acknowledged 'dead' theories are 100% wrong — Lamarck was dismissed years ago, but today his ideas spur valuable science.

Rose argues that a single *g*-loaded coefficient, such as an IQ score, cannot capture all of intelligence. We agree. However, the limitations of IQ are not a reason to muzzle researchers who believe that, albeit imperfect, it nevertheless provides important predictive information.

Rose argues that the dubious biological underpinnings of race render scientific analyses meaningless. Although we agree that the terms white, black, Caucasian or Asian do not define distinct races, this misses the point: self-defined race often provides the largest coefficient in models of achievement. Even if race is a proxy for a chain of social and environmental events that are not genetic, this does not minimize the scientific value of studying it. No one who self-defines as black will deny that this identity is omnipresent in their own and others' consciousness, influencing social interactions and stirring identity threat. Most governments require reporting of racial gaps; numerous studies target narrowing them. Such factors justify studying race, even if it is entirely socially constructed.

The same goes for gender, which — as Rose points out — is highly influenced by socialization but nonetheless continues to offer important insights as a category. Empirical research reveals cognitive differences between the sexes in 4-month-old babies. Even if one argues that these infants' environment has already caused sex differences, this does not gainsay their empirical basis nor their scientific validity.

No one endorses pseudoscience, but Rose has prejudged what qualifies.

Wendy M. Williams, Stephen J. Ceci
Department of Human Development, Cornell University, Ithaca, New York, USA
e-mail: wmw5@cornell.edu

Further contributions are welcome online at http://tinyurl.com/askwhp.

Human rights cannot cover cells that were never in the womb

SIR — In his contribution to the Commentary 'Your inbox, Mr President', in which leading scientists provided advice to Barack Obama (*Nature* **457,** 261; 2009), George Daley writes of the strong indication that the Obama administration will remove restrictions on federal funding for human embryonic stem-cell research. Although Daley recognizes the vital importance of this innovative research, he also illustrates the potential limitations and legislative or regulatory restrictions.

The controversial and moral question of when human life actually begins is the basis of limitations placed on work such as somatic-cell nuclear transfer or the derivation of new cell lines from embryos discarded from *in vitro* fertilization procedures. On 29 May 2008, the Brazilian Supreme Court approved legislation allowing stem-cell research. One of the arguments for this approval is that embryos that have already been frozen would never be implanted in the uterus.

Although many people affirm that human life begins when the sperm is impregnated into the egg, either naturally or artificially, they ignore the role of the uterus. Human status is attributed by some to 5-day-old embryos, which consequently would receive the same protection rights as any person. Do frozen and/or inviable embryos, which could never be implanted in a maternal uterus, deserve this status? To begin the process of forming a living human, there must be connection of an embryo to a mother's epithelial cells. This requires the interaction of the embryo with the endometrium (the wall of the uterus).

The Brazilian constitution states that the civil rights of a person begin at the moment of live birth. The law extends these neonate rights from the moment of conception — but only if the baby is born alive. A fertilized egg that is not implanted in the uterus is neither a neonate nor a person. According to traditional Roman law: "*Nasciturus pro iam nato habetur, quotiens de commodis eius agitur*", which translates as "The unborn child is treated as a child already born in all things respecting its interests". This was presumably conditional on its being born alive. If not, then this protection becomes obsolete.

For a human being to exist and to benefit from civil protection laws, the uterus is a crucial component.

Patricia Pranke Hematology and Stem Cell Laboratory, Federal University of Rio Grande do Sul, and Stem Cell Research Institute, Av Ipiranga, 2752, 304G, 90610-000, Porto Alegre, Brazil
e-mail: patriciapranke@ufrgs.br
João Carlos Silveiro Stem Cell Research Institute and Silveiro Advogados, Av Dom Pedro II, 1240, 90550-141, Porto Alegre, Brazil

© 2009 Macmillan Publishers Limited. All rights reserved

COMMENTARY

A smarter way to combat hunger

Traditional approaches to supplying food are an inefficient 'band aid', says **Pedro A. Sanchez**. New evidence shows that helping farmers to help themselves is more effective and would be six times cheaper.

After decades of progress in the fight to vanquish world hunger, the number of undernourished people is growing again. Estimates from the Food and Agriculture Organization of the United Nations suggest that 963 million people[1] in poor countries are chronically or acutely hungry — up 109 million from 2004 estimates[2]. The underlying causes — changes in food and energy prices[3] — have been exacerbated by the financial crisis and obsolete development policies.

Policies should shift from prioritizing food aid to providing poor farmers with access to training, markets and to farm inputs such as fertilizer and improved seed. In addition to being cheaper, such investments allow farmers to grow food to feed themselves, to sell the surplus and to diversify into high-value crops, livestock and tree products. This creates a sustainable exit from the poverty trap, thereby decreasing the requirement for aid. Although marginal populations, or those affected by disasters, will still require assistance, procuring this food from within developing countries provides a cheaper alternative than shipping it from abroad.

The predominant policies to tackle hunger epitomize a 'band-aid' approach — quick fixes that fail to address the causes of hunger. In 2006, the United States spent US$1.2 billion in food aid for Africa, but only $60 million on agricultural development there[4]. The international response has generally been similar. But according to estimates from 2004, only 10% of those who are hungry in poor countries are acutely hungry — those facing famine caused by wars, natural disasters or sheer destitution. The other 90% are chronically hungry, leading to malnutrition that compromises immune systems and contributes to the prevalence of diarrhoea, malaria and other diseases that result in high child mortality[2]. Most of those who are chronically hungry live in rural farm households in Africa and South Asia.

Food aid fails to provide a sustainable solution to hunger and poverty and it is comparatively expensive. It costs $812 to deliver one tonne of maize as US food aid to a distribution point in Africa[5]. As part of the Millennium Villages project, which I co-direct, smallholder farmers (those who farm 0.1–5 hectares) in hunger hot spots across Africa were provided with access to fertilizers, improved seed, technical support and markets. As a result, maize yields more than doubled — from 1.7 to 4.1 tonnes per hectare[6]. And following a national 'smart' subsidy programme for fertilizer and hybrid seed in Malawi, average maize yields increased from 0.8 to 2.0 tonnes per hectare in two years[7].

The fertilizer and improved seed required to produce an additional tonne of maize grain by Millennium Village farmers cost an average of $135 at April 2008 prices[6], six times less than through food aid. Purchasing that same tonne of maize locally — in an African country or a neighbouring one — costs approximately $320 (ref. 5). If farmers in Africa raise their average cereal yields to 3 tonnes per hectare, the additional 200 million tonnes grown in the 100 million hectares of smallholder crop land will more than compensate for the 3.2 million tonnes of food aid[8].



P. SANCHEZ

Although estimates of efficiency vary, they indicate a major leap for development assistance. Shifting 50% of the current US food-aid budget to 'smart' subsidies or credit could help millions supply their own food and meet much of the aid demand. Such a move would be budget neutral.

Even buying food locally represents an important step away from the inefficient food-aid approach. Some institutions have already begun to change their methods. In 2007, CARE International, a leading relief organization headquartered in Atlanta, Georgia, announced that it would stop monetizing food aid (selling some of the food to fund their operations), essentially losing $46 million a year. Also in 2007, the World Food Programme procured 43% of the 2 million tonnes of food required for its Africa relief operations from farmers in Africa at an average cost of $280 per tonne — compared with the average cost of $436 for purchases elsewhere[9]. The new Purchase for Progress programme, launched in September 2008 and funded by the Bill & Melinda Gates Foundation further empowers the World Food Programme to purchase food from African farmers.

Most importantly, the UN secretary general Ban Ki-moon is leading the development of a coordination mechanism for large-scale financial support for poor countries seeking to provide farm investment. The Spanish government has pledged €1 billion (US$1.3 billion) over five years for this effort, which should begin this year, and the European parliament has promised a similar amount. With more programmes aimed at merging food aid with reliable farming investment, the numbers of those who are chronically hungry should begin to fall. ■

Pedro A. Sanchez is senior research scholar and director of tropical agriculture at the Earth Institute at Columbia University, 61 Route 9W, Palisades, New York 10964, USA.
e-mail: psanchez@ei.columbia.edu

1. Diouf, J. Speech at *High-Level Meeting on Food Security for All* Madrid, 26–27 January 2009. Available at www.ransa2009.org/docs/docs/speech_DG_FAO_ransa2009.doc.pdf
2. Sanchez, P. A. & Swaminathan, M. S. *Science* **307,** 357–359 (2005).
3. von Braun, J. *Nature* **456,** 701 (2008).
4. The Chicago Initiative on Global Agricultural Development *Renewing American Leadership in the Fight Against Global Hunger and Poverty* (Chicago Council on Global Affairs, 2009). Available at http://www.thechicagocouncil.org/globalagdevelopment/pdf/gadp_final_report.pdf
5. Garrett, L. A. *Food Failures and Futures* Maurice R. Greenberg Center for Geoeconomic Studies Working Paper (Council on Foreign Relations, 2008).
6. Sanchez, P. A., Denning, G. L. & Nziguheba, G. *Food Security* **1,** 37–44 (2009).
7. Denning, G. *et al. PLoS Biol.* **7,** e1000023 (2009).
8. http://one.wfp.org/interfais/2008/tables/Table10.pdf
9. Jury, A. *New Roles for Food Assistance: How Can Food Aid Support Agricultural Growth and Productive Safety Nets in Africa?* Presentation at the World Food Prize Symp., 22 October 2008 (World Food Programme, 2008).

Join the debate at http://tinyurl.com/cy3xc6.

© 2009 Macmillan Publishers Limited. All rights reserved

BOOKS & ARTS

Shining light upon light

Two science histories dissect the transfer of knowledge between the Greco–Islamic and European civilizations, and put right the impression that the flow was one way, explains **Yasmin Khan**.

It has been widely accepted that the Islamic civilization had merely a bridging role in preserving the wealth of inherited ancient Greek knowledge ready for future consumption by the West. This pervasive belief, now known to be a damaging distortion of history, is explored in two new books.

In *Aladdin's Lamp*, John Freely, a seasoned academic and historian, writes a captivating account of the transfer of scientific ideas between these civilizations. Interlacing historical events with finesse, his story has a nostalgic quality that makes for escapism but falls short of convincing. At first glance, we assume that Freely will offer us an exposé of the central part the Islamic world played in the pursuit of science, and the key contributions it made. Instead, it quickly transpires that Freely's handling of Islamic discoveries could be construed as damning with faint praise in comparison with his treatment of Greek knowledge.

Freely introduces his book by declaring that "Modern Science traces its origins back to ancient Greece", arousing suspicion that his motive is to venerate the ancient Greeks as the progenitors of scientific ideas, and to suggest that later civilizations should be viewed as being in their shadow. By the end of the book it becomes apparent that this suspicion is founded. Yet Freely's thesis raises the question of whether the emergence of modern science, as practised today, really was spearheaded by the ancient Greeks. Old-school historians were adamant that the scientific revolution emerged in the sixteenth and seventeenth centuries, when the original sources of classical Greek thinking were 'rediscovered' by Europe.



From the ninth century, Islamic astronomers used translated Greek texts to develop theories that enlightened scientists in Europe.

G. DAGLI ORTI/UNIV. LIBRARY ISTANBUL/THE ART ARCHIVE

Aladdin's Lamp: How Greek Science Came to Europe Through the Islamic World
by John Freely
Alfred Knopf: 2009. 320 pp. $27.95

Science and Islam: A History
by Ehsan Masood
Icon Books: 2009. 256 pp. £14.99

Others, such as Ehsan Masood, beg to differ. In *Science and Islam* — the accompanying book to a BBC television and radio series that focuses on science in the Muslim world — he shows that the information flowed in two directions. Through a translation movement that began in the early ninth century, the Islamic world built extensively on Greek ideas, as well as on knowledge from other civilizations, to develop new theories. A golden age for the Islamic civilization, this prolific period spanned more than 800 years. Scientific, technological and engineering endeavour was cultivated to such an extent that it attracted interest in Europe, which was supposedly languishing in the Dark Ages. We now know that the notion of a Dark Age was a myth, constructed to make the Renaissance seem all the more spectacular.

Similarly, George Saliba's iconoclastic *Islamic Science and the Making of the European Renaissance* (MIT Press, 2007) argues that much Islamic scientific thought was highly original. He vehemently maintains that the astronomers who worked under Islam, such as Nasir al-Din al-Tusi and ibn al-Shatir, overhauled Ptolemy's Greek *Almagest* through commentary, correction, objection and ultimately reformulation. The planetary theories that were developed by Islamic astronomers, argues Saliba, would later have a major influence on Nicolaus Copernicus.

Remarkably, besides the Greeks, Freely overlooks cross-cultural influences that were infused into Islamic scholarship. For example, when the art of paper-making was learnt from the Chinese in the eighth century, its growth into an industry in Baghdad further facilitated the translation movement, transforming learning forever. Caliph Harun al-Rashid was among the Muslim rulers who promoted the mass translation of Greek texts into Arabic; al-Rashid's Baghdad is the setting for the fictional story of Aladdin from the *Thousand and One Nights*. Aladdin's ever-glowing magical lamp is portrayed by Freely as a symbolic baton of knowledge that passed from the Greeks to the Arabs and then onto the West. This metaphorical use of light as a beacon of knowledge is in line with the tainted accumulative view of history: when the lights flickered out in one dying empire, a new erudite civilization was ready to stoke the flames, thus shedding light upon light. By contrast, Masood observes that the exchanges of knowledge took place within an interdependent framework; a far-reaching and complex web that was more interconnected than Freely suggests.

Freely does acknowledge a masterpiece produced by Ibn al-Haytham (known in the West as Alhazen) — the *Book of Optics*, which is considered one of the most influential works produced in Islamic science, representing a definitive advance beyond the achievements of the ancient Greeks in their study of light. Masood elaborates further, asserting that al-Haytham pioneered a progenitor to the modern scientific method back in the eleventh century. Al-Haytham's investigations were based on experimental rather than abstract evidence, and his experiments were systematic and repeatable, enabling him to establish empirical proof of the intromission theory of light — that vision is the result of light from objects entering the eye. Two centuries later,

© 2009 Macmillan Publishers Limited. All rights reserved

al-Haytham's work had a profound influence on Roger Bacon, the Arabic-speaking Oxford scholar who was a protégé of the pupils of the Spanish Moors; Bacon quotes al-Haytham throughout his *Opus Majus*. It was not until the seventeenth century that a different historical figure, Francis Bacon, would champion the modern scientific method as part of the scientific revolution.

Masood asserts that the eighth-century chemist Jabir ibn Hayyan produced texts rich with laboratory techniques and experimental methods that are familiar in modern-day chemistry. Jabir came up with a framework for classifying materials and devised an array of chemical processes including sublimation, liquefaction, crystallization and distillation, and discovered both alkalis and strong acids such as sulphuric and hydrochloric acid. Subsequently, al-Razi (also known as Rhazes) built on Jabir's work and surpassed him; his practical work reinforced the need for proof by experimentation, enabling him to contribute to the development of pharmacology.

Among other European scholars who Masood explains were profoundly influenced and inspired by Muslim polymaths were Leonardo da Vinci and Johannes Kepler. Masood highlights some of the discoveries, scientific instruments and inventions made by Muslim scholars in the realms of medicine, surgery, chemistry, physics and astronomy.



Nasir al-Din al-Tusi's early work on planetary orbits may have later inspired Nicolaus Copernicus.

BRITISH LIBRARY

A science journalist and editor at *Nature*, Masood manages to address the subject in a refreshingly non-partisan fashion, presenting a spectrum of arguments from the most relevant academics and historians. He is bold enough to present counter-arguments from adversarial sources, to the extent that he seems polemical. The result is a successful compilation of a complex and contentious history that is conveyed with simplicity. Because *Science and Islam* is aimed at a popular audience, many intricate details are missing. Yet it is the first concise book on the topic suitable for the non-expert.

Both books are opportune and contribute to the long-overdue popularization of the multicultural history of science. No doubt a flurry of similar books will shortly appear, especially given the current political climate coupled with the underpinning role that science has in modern society and the possibilities for development it offers in reviving the Middle East. Yet what is still needed is an updated popular historiography that can span the full breadth of world history and position the outputs of Islamic science into a wider context. That is worth waiting for. ■

Yasmin Khan is Curator Team Manager at the Science Museum, London SW7 2DD, UK, part of the National Museum of Science and Industry.
e-mail: yasmin.khan@sciencemuseum.org.uk

Myth of the missing mothers

Motherhood, the Elephant in the Laboratory
Edited by Emily Monosson
Cornell Univ. Press: 2008.
232 pp. $25, £13.95

Confront a scientist with data and they will suggest a hypothesis to explain them. When confronted with the dearth of women in the scientific profession, scientists often try to connect gender — female — with a physiological reason for absence — motherhood. Family, they say, must be the reason that there are so few women in the sciences, as the years when women start families coincide with those when young scientists are building academic careers.

This hypothesis does contain a kernel of truth: the academic profession, which evolved from medieval monasteries, was not designed with a dual-career family in mind. The modern workplace is still evolving, far too slowly, towards a gender-equal world, where both parents have busy lives and both can pursue their scientific passions.

Motivated by her own struggles to balance motherhood and work, toxicologist Emily Monosson collects in *Motherhood, the Elephant in the Laboratory* the parenting experiences of nearly three dozen mothers and scientists, two of whom are her daughters. She groups these viewpoints by era, starting with women who came of age scientifically in the 1970s and going up to the present day. One sees the gradual changes that have resulted from, and helped to support, a steadily increasing presence of women in science.

But progress is not steady enough. Those of us who raised families in the 1990s pushed to replace 'disability leave' with proper 'maternity leave' and to create rooms where mothers could express breast milk. We tried to get on-site day care, although we were rarely successful. We thought these accommodations were a good start and a done deal. Yet, as the disturbing stories of women who got their PhDs in the past decade make clear, university women still confront ignorance and inflexibility. Many practices that should have vanished — lack of maternity leave or the inability of employers to cope with pregnant employees — are still with us.

> "It is harder to raise a family as a supermarket employee than as a professor of physics."

Evolutionary biologist Gina Wesley-Hunt (PhD 2003), tells us she "was fired for getting pregnant". Katherine Douglass, a physician and Monosson's daughter, describes the difficulty of finding a romantic partner who would view her as an equal. A young nursing mother, Deborah Duffy (PhD, Psychology, 2001), recalls an all-day interview for a faculty position without the requested breaks for pumping breast milk. "As I sat there doing my best impression of a dairy cow," she writes, "I recall thinking, 'What a bizarre experience.'" My fervent hope is that the experiences described in *Motherhood* become less bizarre — not because female science professors do not have children but because they do, thus helping their colleagues become familiar with the process.

Monosson's book sets out the variety of paths possible for scientist mothers, but it contains few stories about women who have successfully combined traditional careers as science professors with traditional families. I know many such women. I am one, as is Stefi Baum — astronomy professor, mother of four, an old friend and a contributor to the book. It is unfortunate if young women approach their futures convinced that they have to make radical compromises to combine family and

© 2009 Macmillan Publishers Limited. All rights reserved

science. The more young women who push forwards assuming, as young men routinely assume, that it is possible to 'have it all', the sooner that will be possible for everyone. It would be a shame if, after reading this book, young women got the opposite message.

More disturbing is the implication that in the absence of motherhood, women in academic science would have untroubled careers. This is naive. Evidence shows that female scientists without children do not fare better than those with children who remain full-time in the workforce. Neither advance as steadily as their male counterparts, with or without children. Some explanation other than family must be the reason for the slow advancement of women.

Furthermore, countries with enviable support systems for parental leave and childcare — of the kind rightly advocated by Monosson — have astonishingly low participation of women in science. For example, in Denmark, only around 10% of full professors of science, and 2% of physics professors, are women. The numbers are higher in the United States, even with its less generous childcare policies. Neither does physiology and family explain why the fraction of scientists who are women varies widely from one country to another, even when social support systems are similar; the percentage of female scientists is much higher in France and Italy than in the United Kingdom or Germany, for example. Women's participation in science also varies widely from field to field: more than 50% of graduating MDs in US medical schools are women, whereas in physics and engineering, less than 20% of PhDs go to women.

Outside our universities, many women with young children have full-time jobs — 70% in the United States — and few of those jobs are as flexible or as well paid as jobs in academic science. There is no question that it is harder to raise a family as a supermarket employee than as a professor of physics, so why do academics seize on family as the explanation for the absence of women? The dominant obstacles to women in science — persistent, unexamined bias and lack of mentoring — are described clearly in *Beyond Bias and Barriers* (National Academies Press, 2007) and in *Academeology*, by the pseudonymous blogger Female Science Professor (LuLu, 2008; see *Nature* **456,** 445; 2008). Young women trying to figure out the road to success in science might be better served reading those books rather than *Motherhood*. ■

Meg Urry is Israel Munson professor of physics and astronomy and head of the Department of Physics at Yale University, New Haven, Connecticut 06520-8120, USA.
e-mail: meg.urry@yale.edu



Each year, forests release 30% more carbon than is emitted by fossil-fuel burning.

G. LUDWIG/PANOS

Storing carbon in forests

Climate Change and Forests: Emerging Policy and Market Opportunities
Edited by Charlotte Streck, Robert O'Sullivan, Toby Janson-Smith and Richard G. Tarasofsky
Chatham House: 2008. 360 pp. £39.99

Planting, conserving and managing forests are key for attaining long-term mitigation of greenhouse gases, yet the land-use and forestry sector barely features in the Kyoto Protocol. In *Climate Change and Forests*, some 50 experts analyse the forestry-based discussions within climate-change negotiations, and offer technical and political reasons for why the Kyoto Protocol handles forestry in such a cumbersome way.

The authors argue that in the run-up to the 1997 Kyoto conference, the forestry sector's complexities were ignored in negotiations because knowledge about using forests as carbon sinks was limited at the time. This failure to resolve the contributions of forests has meant that rules formulated since then are not conducive to improving the ability of forests to sequester large amounts of additional carbon. The book provides an excellent historical background and describes the process and logic of negotiations, past and present. The multiple analyses of technical problems associated with carbon sequestration and forests make it valuable to both newcomers and veterans in the field.

Case studies demonstrate that the right incentives or policies to change land usage must vary between countries. In developing nations, for example, legal owners of land where carbon sinks could be enhanced are often absent or difficult to identify. Without knowing which people and actions are genuine candidates for carbon-incentive payments, the market-based instruments that help administer these incentives risk losing their efficiency. Experimentation is necessary and may involve many failures before success is achieved.

Although this collection includes many voices, most belong to a single camp that favours market-based instruments for encouraging changes in land use towards carbon storage. This fails to reflect the diversity of opinion in global discussions, and the book would have benefited from the inclusion of some critics. Also missing is evidence for why market-based instruments are superior to other policy interventions. The authors mention the preconditions for markets to work — property rights, legal and regulatory frameworks — but give only vague arguments for how policies could be engineered to support the use of forests in storing more carbon on meaningful scales.

The greatest opportunities for forest-based mitigation of carbon emissions are in the developing world, and the most efficient window in which to stop deforestation is between now and 2030. However, the book fails to address how carbon markets in developing countries, such as the Democratic Republic of the Congo, could be made to operate on a large scale in an institutional vacuum. It is unclear which sources will generate the tens of billions of dollars per year that are needed, for example, to reduce deforestation by half in the coming decades. The funding issue is exacerbated by the current financial crisis that emerged after the book was written.

Climate Change and Forests emphasizes the difficulty of sequestering carbon in forests, given their vulnerability to disturbances that re-emit carbon, such as forest fires. But the book neglects issues of climate adaptation. Forests will take centuries to adapt to the disruptive processes that accompany climate change, making forest carbon stores vulnerable in the long term.

More carbon is locked in forests than in the atmosphere, and the yearly carbon flux attributed to natural disturbances is 30% higher than current emissions of fossil fuels. Under the United Nations Framework Convention on Climate Change, a multibillion-dollar fund to support climate-adaptation measures is being set up. I hope the authors are planning a new book to analyse how the mechanisms around this fund will interact with and strengthen forest-mitigation instruments in the post-Kyoto era. Like *Climate Change and Forests*, this second volume would be a crucial reference for academics, negotiators and policy-makers, and for the forest business community. ■

Michael Obersteiner is a Forestry Program research scholar at the International Institute for Applied Systems Analysis, Laxenburg, Austria.
e-mail: oberstei@iiasa.ac.at

© 2009 Macmillan Publishers Limited. All rights reserved

Between life and death

Mummies: The Dream of Everlasting Life
South Tyrol Museum of Archaeology, Bolzano, Italy
Until 25 October 2009.
www.iceman.it

The preserved Tyrolean Iceman Ötzi is joined by 70 mummies, many displayed for the first time, in an exhibition that opened this week at the South Tyrol Museum of Archaeology in Bolzano, Italy. *Mummies: The Dream of Everlasting Life* examines the science and culture of mummification, with diverse examples of humans and animals preserved in various ways and drawn from South America and Asia to ancient Egypt.

Curated together with the Reiss-Engelhorn-Museums in Mannheim, Germany, the collection includes 20 mummies that were discovered in that museum's basement during renovations in 2004. Thought to have been lost, they had been hidden there for protection during the Second World War.



The rich clothes of this 500-year-old Inca mummy suggest a high social status.

W. ROSENDAHL/BONNER ALTAMERIKA-SAMMLUNG UNIVERSITÄT, BONN

Ötzi the Iceman, the exhibition's host, was found frozen into a glacier in the Ötzal Alps near the Austria–Italy border in 1991. He died there from an arrow wound to the shoulder around 3300 BC, and is thought to have been a hunter, although his identity is unclear. "Since Ötzi cannot leave the museum," explains Bolzano archaeologist Andreas Putzer, "this is a unique occasion to see these mummies together."

The different stages of mummification in ancient Egypt are revealed in a series of partially preserved bodies. A mummified three-year-old Peruvian child from the fifteenth century represents the first evidence from South America of embalming a human corpse with resin, the same technique that was used in Egypt. An eighteenth-century mummified family found in a Hungarian church and a mummified baron from a cellar in Castle Sommersdorf near Nuremberg, Germany, are more recent European examples. Their bodies were dried using artificial air-conditioning systems. "We want to explode the myth that mummies are only from Egypt," says curator Wilfried Rosendahl from the Reiss-Engelhorn-Museums.

A range of mummified animals on display include the only known animal mummy found in a bog — a domesticated dog from the sixteenth century, found near Münster in Germany — and naturally preserved examples, including dinosaurs, from varied environments such as deserts, bogs and caves, or under ice or salt.

Related burial offerings and scientific information are shown beside the mummies. Anthropology, genetics, pathology, toxicology and medical studies provide clues about their living conditions, dietary habits and diseases.

Mummification continues to this day as an alternative to burial or cremation. Methods range from those used in ancient Egypt to cryogenic freezing of the brain, with the hope that one day it will be possible to grow a new body from brain cells by means of DNA replication. "Mummies still trigger a deep interest because they represent something between life and death," says Rosendahl. ■

Marta Paterlini is a journalist based in Stockholm.
e-mail: martapaterlini@nasw.org

A gallery of chimaeric curiosities

Corpus Extremus (LIFE+)
Exit Art, 475 Tenth Avenue, New York
Until 18 April 2009.
www.exitart.org

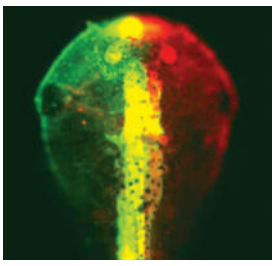
When images of the Vacanti mouse, the mouse with a cartilage ear growing on its back, were released in 1997, it seemed that a surrealist project had come alive. Although the experiment was a step towards using living material for tissue engineering, the surprising and provocative juxtaposition of body parts was also a culturally evocative artefact — an art piece.

The exhibition *Corpus Extremus (LIFE+)*, at New York's Exit Art until 18 April, examines the effect of biotechnology on our notions of life and death. Diverse works include living matter, and reflect on the creation of new organisms — hybrids, cyborgs and extended human bodies.

A cabinet of chimaeric curiosities explores the taxonomical crisis that is presented by new life forms in *NoArk II*, by Oron Catts and Ionat Zurr of SymbioticA, the biological art centre at the University of Western Australia near Perth. They display masses of living cells taken from different organisms alongside preserved museum specimens. Skin and blood cells are used as a screen on which movies are projected in another installation by the BioKino collaboration, a reference to early science films.

The Mirror of Faith, an installation by the Ultrafuturo group, is an attempt to create "self-transcendent" bacteria that carry *VMAT2*, the gene controversially claimed by some to cause higher consciousness and spirituality in humans.

Other works embrace robotics. A culmination of six years of research by artists Guy Ben-Ary and Philip Gamblen, *Silent Barrage* evokes the human machine and futuristic communication interfaces. A robotic body made of 36 connected parts is controlled by electrical signals in a network of rat neurons growing in a remote Petri dish in Steve Potter's lab at the Georgia Institute of Technology in Atlanta.



A fluorescing tadpole reveals its previously unseen developmental secrets.

D. BULATOV, J. LABAS & K. LUKJANOV

Ironic, critical and Utopian, the exhibition raises issues about artificially cultured cells, the presumed uniqueness of human life, our anthropocentric hierarchies and the possibility of engineered and commodified perfection. ■

Giovanni Frazzetto is a Society in Science Branco Weiss Fellow at the BIOS Centre, London School of Economics, Houghton Street, London WC2A 2AE, UK, and is at the European Molecular Biology Laboratory, 00015 Monterotondo, Italy.
e-mail: g.frazzetto@lse.ac.uk

© 2009 Macmillan Publishers Limited. All rights reserved

NEWS & VIEWS

PALAEOANTHROPOLOGY

Asian *Homo erectus* converges in time

Russell L. Ciochon and E. Arthur Bettis III

Re-evaluation of the age of Zhoukoudian, a prominent site of *Homo erectus* occupation in China, prompts a rethink of the species' distribution in both the temperate north and the equatorial south of east Asia.

The Zhoukoudian cave system, discovered in 1918 near the outskirts of modern-day Beijing, is one of the prize sites in palaeoanthropology (Fig. 1). It has produced a long sequence of well-preserved *Homo erectus* fossils and stone tools, unique in northeast Asia, that have been dated variously to between 0.4 million and 0.6 million years (Myr) ago. In their report on page 198 of this issue, Shen and colleagues[1] push back the age of Zhoukoudian's lowest occupation level to 0.78 Myr ago. The new age estimate is consistent with the cold-climate fauna found within the locality's lowest levels, and it also correlates with revisions in the regional chronology of *H. erectus* sites[2–6]. Overall, the recent studies suggest that the onset of glacial cycles brought the open-land habitats in which *H. erectus* thrived.

The first fossils of *H. erectus* were found at Trinil in Java, Indonesia, in 1892 by Eugène Dubois. It was Dubois who coined the name '*erectus*', based on a thigh bone that indicated upright posture[7]. *Homo erectus* had a heavily armoured and thickened skull, with an adult brain size in the 650–1,250-cm^3 range[8] — the equivalent value for modern humans (*H. sapiens*) is 1,100–1,800 cm^3. *Homo erectus* stood 145–180 cm tall[9], walked fully upright with a modern-like human footprint[10], and used stone tools. The species is easily distinguished from *H. sapiens* by its distinctive torso, which was much more barrel-shaped and larger in volume[11].

Fossil evidence now points to an origin of *H. erectus* about 2.0 Myr ago in equatorial Africa, followed by dispersal northwards and eastwards into Eurasia[8]. Incontestable evidence of that dispersal comes from the site of Dmanisi (Transcaucasia, Republic of Georgia) at 1.75 Myr ago and from southeast Asia (Java) by 1.6 Myr ago. *Homo erectus* survived to as late as 0.35 Myr ago in continental east Asia, and isolated populations persisted longer on the Indonesian archipelago. One group on Java, commonly known as 'Solo Man', may have survived until 50,000 years ago[12].

As far as east Asia is concerned, Zhoukoudian has yielded 17,000 stone artefacts and fossil evidence for more than 50 *H. erectus* individuals[8], and so provides a singular window on the life of *H. erectus* in the region's temperate north. Over the past several decades, discoveries in the surrounding basins and plains have broadened our knowledge, yet Zhoukoudian remained a temporal and environmental outlier. Surrounding Zhoukoudian are the open basins and plains comprising the *H. erectus* homeland at 40° N latitude. To the northwest, the lake-bed sediments of Hebei Province hold scattered open-air occupation sites (Maliang, Cenjiawan, Donggutuo, Xiaochangliang). To the southwest, Shanxi Province has yielded cranial remains of *H. erectus* at Gongwangling (Fig. 2, overleaf). Although these areas never saw glacial ice, glacial periods brought cooling and drying conditions. Open habitats, exemplified by grasslands and mixed steppes, expanded during glacial periods. Such environments favoured large grazing animals, with *H. erectus* being one of several large predators.



R. L. CIOCHON, UNIV. IOWA

Figure 1 | Prize site — the western wall of the main excavation at Zhoukoudian. This is all that remains of the massive, decades-long excavation that produced the fossils of 50 *Homo erectus* individuals and thousands of stone artefacts.

Magnetostratigraphy has been the primary dating technique for the lake and loess (wind-blown silt) deposits that harbour China's *H. erectus* record[2–6]. With this approach, one seeks evidence for Earth's known magnetic reversals preserved within a local sedimentary sequence; the age of a buried fossil or artefact horizon is determined by its location in the sequence. Early magnetic analyses produced impressively old ages of 1.7–1.4 Myr ago for Hebei's lake sites[5,6]. At nearly three times older than the previous age estimates for Zhoukoudian, it has been difficult to link the lake sites with the cave.

New work, however, is revealing the problems of open-air sedimentation rates, a crucial variable in magnetostratigraphy[2]. Revised ages have been provided for the

© 2009 Macmillan Publishers Limited. All rights reserved

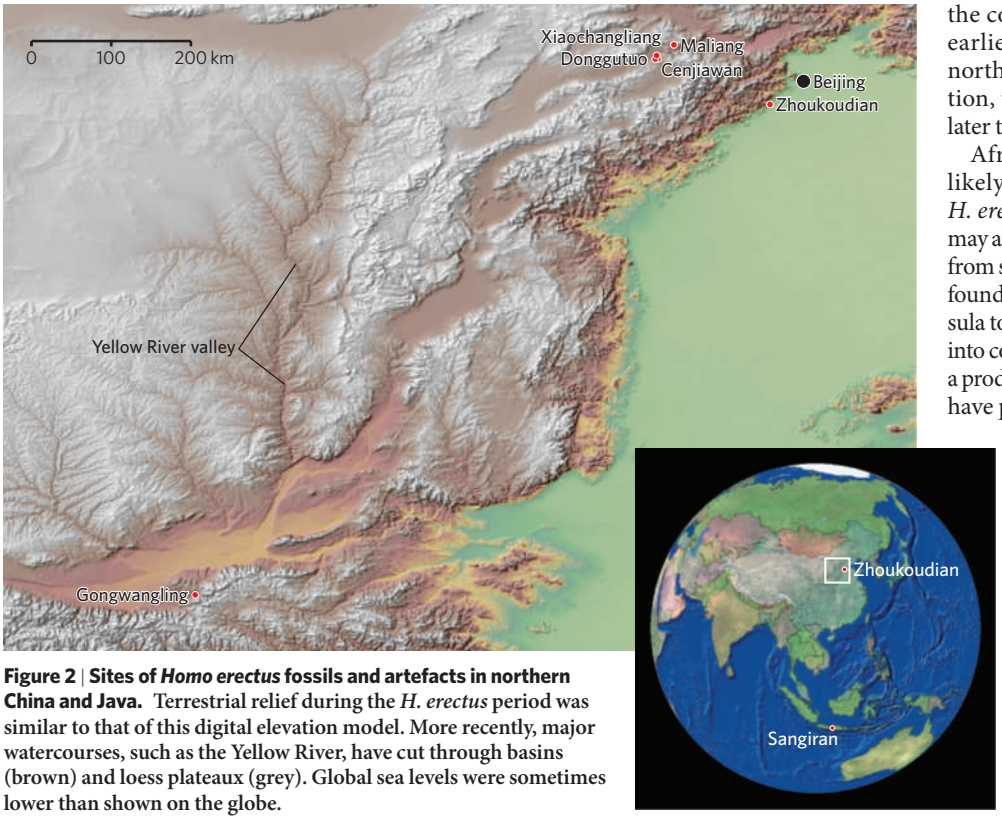


Figure 2 | Sites of *Homo erectus* fossils and artefacts in northern China and Java. Terrestrial relief during the *H. erectus* period was similar to that of this digital elevation model. More recently, major watercourses, such as the Yellow River, have cut through basins (brown) and loess plateaux (grey). Global sea levels were sometimes lower than shown on the globe.

Hebei sites of Maliang (0.78 Myr)[4], Cenjiawan (1.1 Myr)[3], Donggutuo (1.2 Myr)[2] and Xiaochangliang (1.26 Myr)[2], as well as for the Shanxi site of Gongwangling (1.22 Myr)[2]. With Zhoukoudian now estimated to be older, as reported by Shen *et al.*[1] using an approach called cosmogenic $^{26}Al/^{10}Be$ burial dating, and the open-air sites to be younger, it seems that the early human occupation of northeast Asia began around 1.3 Myr ago. That occupation continued to at least 0.4 Myr ago based on the youngest occurrence of *H. erectus* at Zhoukoudian[1].

Still at issue is how *H. erectus* dealt with the colder and drier glacial climates that prevailed at various times during the Pleistocene (1.8–0.01 Myr ago). The loess and palaeosol (ancient soil) sequence in Shanxi Province provides a record of at least ten glacial–interglacial cycles during the period of *H. erectus* occupation[13,14]. Comparing the ages of glacial (loess) and interglacial (palaeosol) periods with those of the *H. erectus* sites suggests that northeast China was occupied in both glacial and interglacial periods. During the early Pleistocene, before about 0.9 Myr ago, the cycles were shorter and less extreme. These conditions may therefore have enhanced the mammal-rich open environments on which *H. erectus* thrived, rather than bringing hardship, as is often thought. Once settled, *H. erectus* adjusted to the lengthening and deepening cycles.

To find anything analogous to Zhoukoudian and its surrounding basins, one must travel 5,000 km south to central Java. At 7° S latitude, the early Pleistocene lake and river deposits at Sangiran have yielded fossils of nearly 80 *H. erectus* individuals and several occupation sites. The *H. erectus* presence at Sangiran has a firm basal age of 1.6–1.5 Myr ago, and begins during an early Pleistocene glacial cycle[15]. A plausible view of events is that low sea level during this glacial period linked some of today's Indonesian archipelago with mainland southeast Asia. A previously mainland fauna, including *H. erectus*, made its way southwards, and thrived in the open woodland and savanna during the relatively dry glacial periods[16]. Interglacials brought higher sea levels, which submerged lowlands and increased moisture, shrinking open habitats. In Java, *H. erectus* persisted through these cycles until the Late Pleistocene[12].

The temperate and equatorial populations were probably not connected. One factor may have been the existence of the subtropical primal forest in southern China and Indochina inhabited by the *Stegodon–Ailuropoda* fauna, which included the giant panda, gibbons, orangutans and the giant extinct ape *Gigantopithecus*[17]. The forest and its fauna may have formed an impenetrable ecological barrier between the two *H. erectus* populations dwelling in open habitats. The picture emerges of separate *H. erectus* populations existing outside a vast and ecologically unattractive forest. Parallel foraging niches existed in open and edge environments in both the temperate and equatorial environments.

Palaeoanthropologists have assumed that east Asian *H. erectus* originated with one group that dispersed just once from a single source in Africa. It may be time to rethink. As evidence for the ability of *H. erectus* to penetrate the subtropical primal forest falls away, so does the connection between Zhoukoudian and earlier Sangiran. If Zhoukoudian and the northern open-air sites represent a population, that population seems to have arrived later than originally thought.

Africa and southwest Eurasia are equally likely as the areas from which east Asian *H. erectus* originated. The two Asian groups may always have been independent and derived from separate dispersals. One group could have found a coastal route across the Arabian Peninsula to the southern flanks of the Himalaya and into coastal southeast Asia[8]; Sangiran represents a product of this dispersal. Another group could have passed through central Asia and southern Mongolia to enter the temperate basins and plains of northeast Asia, Zhoukoudian being a result.

Finally, $^{26}Al/^{10}Be$ burial dating — the technique applied by Shen *et al.*[1] — is of considerable interest in its own right. It capitalizes on the formation of unstable isotopes, ^{26}Al and ^{10}Be, in quartz grains exposed to cosmic radiation at the land surface, and their differential decay in deeply buried sediments shielded from cosmic radiation. In caves, where surface sediment rapidly becomes deeply buried, burial dating provides a cross-check on age estimates derived by magnetostratigraphy. The approach also serves as a dating option in situations where magnetostratigraphy is problematic.

Early Pleistocene lake sediments in southwest China have been recently dated using the new technique[18]. This study used burial ages to constrain the age of fault movements that influenced the evolution of the middle Yangtze River. Likewise, $^{26}Al/^{10}Be$ burial dating will advance our understanding of *H. erectus* throughout its range by providing a new way of dating the Hubei Basin lake sites, as well as other *H. erectus* sites throughout east Asia. ■

Russell L. Ciochon is in the Department of Anthropology, and E. Arthur Bettis III is in the Department of Geoscience, University of Iowa, Iowa City, Iowa 52242, USA.
e-mails: russell-ciochon@uiowa.edu; art-bettis@uiowa.edu

1. Shen, G., Gao, X., Gao, B. & Granger, D. E. *Nature* **458,** 198–200 (2009).
2. Li, H. *et al. Quat. Res.* **69,** 250–262 (2008).
3. Wang, H. *et al. Sci. China (D)* **49,** 295–303 (2006).
4. Wang, H. *et al. Quat. Res.* **64,** 1–11 (2005).
5. Zhu, R. X. *et al. Nature* **431,** 559–562 (2004).
6. Zhu, R. X. *et al. Nature* **413,** 413–417 (2001).
7. Shipman, P. & Storm, P. *Evol. Anthropol.* **11,** 108–116 (2002).
8. Boaz, N. T. & Ciochon, R. L. *Dragon Bone Hill: An Ice Age Saga of* Homo erectus (Oxford Univ. Press, 2004).
9. Lordkipanidze, D. *et al. Nature* **449,** 305–310 (2007).
10. Bennett, M. R. *et al. Science* **323,** 1197–1201 (2009).
11. Franciscus, R. G. & Churchill, S. E. *J. Hum. Evol.* **42,** 303–356 (2002).
12. Swisher, C. C. III *et al. Science* **274,** 1870–1874 (1996).
13. Zhou, L. & Shackleton, N. *Earth Planet. Sci. Lett.* **168,** 117–130 (1999).
14. Liu, T. & Ding, Z. *Annu. Rev. Earth Planet. Sci.* **26,** 111–145 (1998).
15. Larick, R. *et al. Proc. Natl Acad. Sci. USA* **98,** 4866–4871 (2001).
16. Bettis, E. A. III *et al. J. Hum. Evol.* **56,** 11–24 (2009).
17. Rink, W. J. *et al. Quat. Res.* **69,** 377–387 (2008).
18. Kong, P. *et al. Earth Planet. Sci. Lett.* **278,** 131–141 (2009).

© 2009 Macmillan Publishers Limited. All rights reserved

CARBON CYCLE

Fickle trends in the ocean

Nicolas Gruber

A model analysis of the uptake of carbon dioxide in the North Atlantic carries with it a cautionary reminder about interpreting what may be short-term trends as signals of long-term climate change.

Several observational studies have suggested that the carbon sink in the North Atlantic Ocean has decreased in recent decades, possibly reflecting the impact of long-term climate change. Writing in *Global Biogeochemical Cycles*, Thomas *et al.*[1] propose that this latter interpretation needs to be viewed with great caution. From their modelling work, they argue that the observed trends instead reflect fluctuations on a decadal timescale that are a response to climate variability in the North Atlantic region.

An underlying question in this line of research is whether the global carbon cycle has already been subject to fundamental alteration as a result of anthropogenic climate change. This question is often asked in the context of concern that the ocean's carbon sinks, and those on land, will not continue to mitigate anthropogenic carbon dioxide emissions as they have in the past[2]. One modelling study[3], for example, has suggested that the oceanic sink, which has removed about 30% of the global anthropogenic emissions over the past 250 years or so[4], might be stalling. Some of the best observational evidence comes from the North Atlantic Ocean, where long-term measurements of the surface ocean's partial pressure of CO_2 (pCO_2) indicate that its carbon uptake from the atmosphere has decreased in recent decades, perhaps owing to climate change[5,6]. Thomas *et al.*[1] challenge this interpretation.

The North Atlantic is the largest ocean sink for atmospheric CO_2 in the Northern Hemisphere, with half of the flux in the North Atlantic being driven by the uptake of anthropogenic CO_2 (ref. 7). The detection of long-term changes in this sink is challenging, however, because the sink varies substantially from year to year. That variation is largely associated with the North Atlantic Oscillation (NAO)[8], which is the dominant mode of climate variability in this region.

The NAO is a large-scale seesaw in atmospheric mass between a subtropical high-pressure system, typically near the Azores, and a subpolar low near Iceland. A positive phase of the NAO — that is, a stronger pressure gradient between these two systems — is associated with more and stronger winter storms crossing the North Atlantic on a more northerly route, causing major anomalies in sea surface temperature, currents and convective activity throughout the North Atlantic (Fig. 1). The NAO is characterized by substantial decadal trends, starting with a period of negative values in the 1960s, a long-term trend towards very positive values that culminated in the early 1990s, and a decreasing trend since.

Using a global model of the ocean carbon cycle, forced with observed atmospheric conditions, Thomas *et al.*[1] investigated how these variations and trends in the NAO have affected the North Atlantic carbon sink since 1979. They find substantial year-to-year changes in the surface-ocean carbon cycle, with an overall tendency for increased uptake in the temperate and subpolar North Atlantic during positive phases of the NAO. This response is driven primarily by large-scale reorganizations of the North Atlantic circulation (Fig. 1).

Thomas and colleagues' comparison of the model-simulated variations in pCO_2 with observational data suggests that the model captures the main characteristics of the oceanic response to NAO variability. This permits them to put analyses of relatively short-term observational trends[5,6] into a longer-term context. They point out that most of the observations were made after the early 1990s, a period during which the NAO changed from very positive to near normal values. Their model indicates that normal to negative phases of the NAO are associated with a lower oceanic uptake of CO_2; so the trend towards lower NAO states since the early 1990s could explain the tendency of the North Atlantic to take up less CO_2 than expected during this period. Thomas *et al.* also speculate that the North Atlantic carbon sink will probably rebound in the coming years, once the present NAO trend reverses.

The message that analyses of trends occurring over short periods need to be interpreted with great caution is clearly an important one — not least because scientists are often tempted to claim the detection of a climate-change signal in the absence of a full characterization of the variability. That said, are their conclusions robust? Or is it nevertheless possible that the carbon cycle in the North Atlantic has already undergone a fundamental shift?

It is clearly too early to draw final conclusions. But it is of interest that an extended analysis[9] of long-term observations indicates that the carbon sink in several regions of the North Atlantic could have been decreasing for several decades already, implying that Thomas and colleagues' analysis of NAO-driven variability is only part of the puzzle.

What about carbon sinks in other parts of the global ocean? With a few notable exceptions[9],

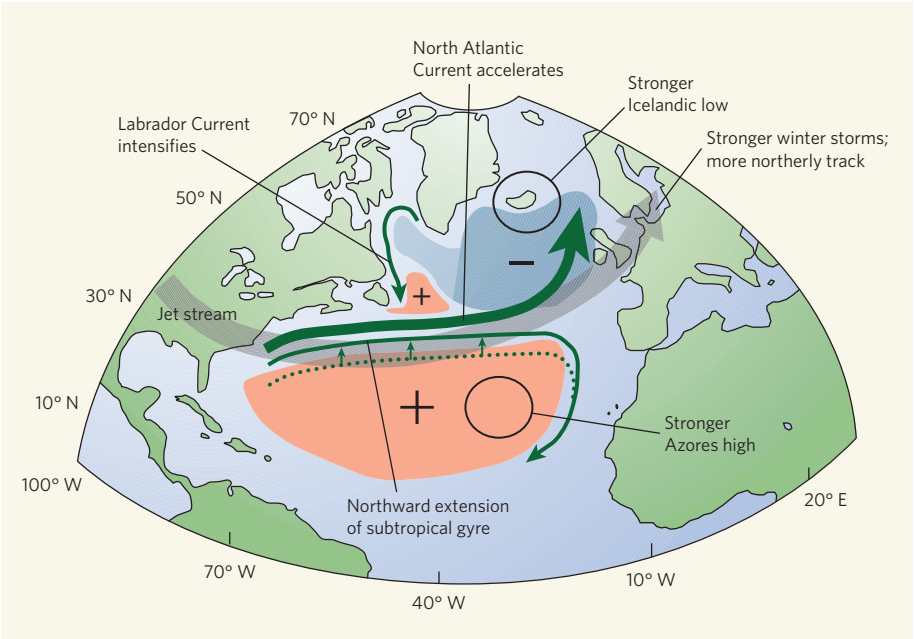


Figure 1 | Impact of a positive phase of the North Atlantic Oscillation. In a positive phase, a stronger Azores high and stronger Icelandic low produce more and stronger winter storms on a more northerly track. As a consequence, the subtropical gyre extends northwards, and the North Atlantic Current accelerates, transporting increased amounts of warm, saline waters with low carbon concentrations northeastwards. This causes an intensified sink (minus sign) in the eastern subpolar North Atlantic, because these waters have the potential to take up a large amount of CO_2 from the atmosphere when they are cooled along their northward journey. At the same time, the Labrador Current intensifies, bringing fresher, colder waters with high carbon concentrations from the Arctic into the subpolar gyre, creating a diminished sink (plus sign) near the coast of Canada. In the subtropical gyre, warm conditions and reduced convective activity also lead to reduced carbon uptake[8]. (Graphic based on the modelling results of Thomas *et al.*[1].)

© 2009 Macmillan Publishers Limited. All rights reserved

observational records are generally too sparse to determine long-term changes in the oceanic carbon sink on a regional basis. One area that has received much attention recently, however, is the Southern Ocean, because it is the strongest regional oceanic sink for anthropogenic CO_2 (ref. 7). Observations and modelling studies[3,9,10] suggest that the oceanic uptake there has slowed considerably relative to the expected growth of the sink in response to the rise in atmospheric CO_2. But the observational evidence is limited, and the modelling studies may be biased[11], so it is not yet possible to determine with sufficient confidence whether or not the Southern Ocean sink is changing in a fundamental manner. Nor is it possible to make such conclusions for any other large-scale oceanic region.

It is essential to resolve this issue. The first task is to ensure that appropriate observational systems are put in place to permit accurate quantification of the oceanic carbon sinks and detect changes reliably. We also need to improve our understanding of the underlying processes so as to better assess how the ocean carbon cycle will behave in the future. As international agreements to limit the further growth of atmospheric CO_2 come closer to reality, accurate determination of the fate of anthropogenic CO_2 emissions is becoming a matter of ever-growing importance. ■

Nicolas Gruber is in the Environmental Physics Group, Institute of Biogeochemistry and Pollutant Dynamics, ETH Zurich, 8092 Zurich, Switzerland.
e-mail: nicolas.gruber@env.ethz.ch

1. Thomas, H. *et al. Glob. Biogeochem. Cycles* doi:10.1029/2007GB003167 (2008).
2. Friedlingstein, P. *et al. J. Clim.* **19,** 3337–3353 (2006).
3. Le Quéré, C. *et al. Science* **316,** 1735–1738 (2007).
4. Sabine, C. L. *et al. Science* **305,** 367–371 (2004).
5. Lefèvre, N. *et al. Geophys. Res. Lett.* doi:10.1029/2003GL018957 (2004).
6. Schuster, U. & Watson, A. J. *J. Geophys. Res.* doi:10.1029/2006JC003941 (2007).
7. Gruber, N. *et al. Glob. Biogeochem. Cycles* doi:10.1029/2008GB003349 (2009).
8. Gruber, N., Keeling, C. D. & Bates, N. R. *Science* **298,** 2374–2378 (2002).
9. Takahashi, T. *et al. Deep Sea Res. II* doi:10.1016/j.dsr2.2008.12.009 (2009).
10. Lovenduski, N. S., Gruber, N. & Doney, S. C. *Glob. Biogeochem. Cycles* doi:10.1029/2007GB003139 (2008).
11. Böning, C. W., Dispert, A., Visbeck, M., Rintoul, S. R. & Schwarzkopf, F. U. *Nature Geosci.* **1,** 864–869 (2008).

NEUROSCIENCE

Up, down, flying around

Ruth Anne Eatock

The Johnston's hearing organ of the fruitfly has newly discovered sensitivities to gravity and wind. As in our inner ear, different sensory signals from this organ travel in parallel to separate zones in the brain.

The hearing organ in the antennae of mosquitoes and flies — first described by Johnston[1] — detects sounds from nearby sources such as the vibrating wing of the courting male fruitfly. In this issue, Kamikouchi *et al.*[2] and Yorozu *et al.*[3] dramatically expand the number of known effective stimuli for Johnston's organ to include gravity and wind. These results elevate Johnston's organ from a near-field hearing sensor to a complex ear in which specialized clusters of neurons have distinct mechanosensory responsibilities.

Johnston's organ is the largest of the chordotonal organs that are located beneath the epidermal layer of the insect skin. Chordotonal organs are assemblies of units each of which consists of a ciliated neuron and several associated support cells[4]; Johnston's organ in the fruitfly contains hundreds of such units. Nearby sounds vibrate feathery structures (aristae) on the antennae, stretching cilia on Johnston's organ neurons (Fig. 1) and thereby exciting them to make sensory signals. The molecular mechanism of transduction is obscure, but sensory-neuron signals can be visualized by measuring increased levels of intracellular calcium ions, which universally accompany neuronal excitation. Kamikouchi *et al.* and Yorozu *et al.* therefore genetically engineered Johnston's organ neurons to express molecules that fluoresce when calcium-ion concentration rises. They then compared the calcium-ion signals evoked by sound, wind, gravity or movement of the aristae, looking at neurons either in Johnston's organ[1] or in the brain's antennal mechanosensory centre[2].

The need to know which way is up is considered so pressing that it has been assumed that flies are sensitive to gravity. These insects lack an obvious statolith — crystal aggregates that weight gravity-sensing receptors in organisms as diverse as plants, octopuses and humans. In the absence of statoliths, chordotonal organs on the flies' legs[5] and Johnston's organ[6] were considered likely candidates for gravity sensing.

Kamikouchi *et al.*[2] (page 165) show that it is Johnston's organ that is responsible for gravity sensing in flies. They find that removing aristae severely decreases the tendency of flies to walk upwards (the direction opposite to the gravity vector). Their test was not designed to rule out minor contributions from neurons outside Johnston's organ to the overall sensing of gravity. Nevertheless, it clearly shows that flies detect a change in their orientation relative to gravity through the effect of gravity on the position of the aristae. The authors' calculations, taking into account an arista's apparent mass and stiffness, suggest that gravitational changes maximally deflect an arista by about one micrometre — a movement 100 times greater than that evoked by sound at the threshold of hearing[7].

A previous study[8], and now that of Yorozu and colleagues[3] (page 201), also used simple antennal manipulations to investigate the role of Johnston's organ in sensitivity to wind. The earlier paper[8] showed that gentle air currents such as those experienced in flight modify flight behaviour, an effect that is abolished by removal of the aristae. Yorozu *et al.* find that, in response to air currents moving at about 2 metres per second, fruitflies 'freeze' in place — a behaviour whimsically dubbed WISL, for wind-induced suppression of locomotion, and which also depends on the presence of the arista. The benefits of hunkering down in the wind may range from a short-term individual advantage of staying close to familiar terrain and food sources to a longer-term group benefit of controlling the dispersal of fly populations.

Whereas courtship song vibrates the aristae at frequencies of tens to hundreds of hertz, wind and gravity stimuli change more slowly and may be sustained. Both teams[2,3] tested the importance of this difference for stimulus discrimination within Johnston's organ by using probes to displace the arista at different frequencies or durations. They find that fast and slow stimuli differentially activate neuronal groups that originate from distinct zones in the Johnston's organ array and target specific zones (A–E) in the antennal mechanosensory centre. Sound-sensitive neurons of zones A and B responded briskly to arista vibrations or to the onset of position changes; when new positions were sustained, however, their response adapted by declining. Wind- and gravity-sensitive neurons of zones C and E, by contrast, responded poorly to vibrations and relatively slowly to position changes, but showed robust, sustained responses to long-lasting stimuli.

These observations indicate that the main difference between sound-sensitive and wind- or gravity-sensitive neurons might be in the speed at which their transduction cascades respond and adapt. This difference could reflect quantitative or qualitative differences in the transduction pathways of the neuronal groups involved. Indeed, Kamikouchi *et al.*[2] identify a qualitative difference, in that only the sound-sensing neurons express a particular stretch-activated ion channel called NompC (ref. 5).

But how do fruitflies distinguish wind from gravity? These two kinds of input are also likely to differ intrinsically from each other in duration and size — although the differences

© 2009 Macmillan Publishers Limited. All rights reserved

J. SCOTT & R. BHATNAGAR, DEPT BIOL. SCI., UNIV. ALBERTA

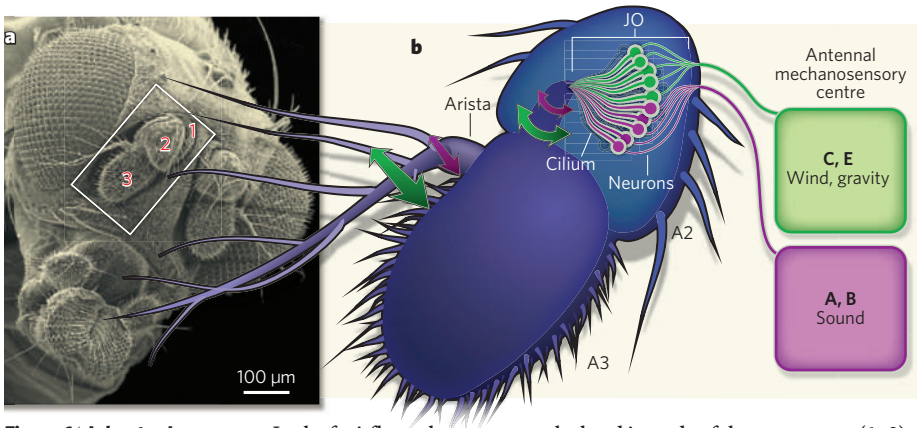


Figure 1 | Johnston's organ. **a**, In the fruitfly, each antenna on the head is made of three segments (1–3). **b**, Johnston's organ (JO) — which consists of hundreds of arrayed neurons, each with a sensory cilium that receives stimuli — is located beneath the epidermis covering segment 2 (A2) and functions by detecting rotation of segment 3 (A3) relative to A2. Near-field sound produces minute high-frequency deflections of the feathery arista connected to A3. The resulting signals project through a distinct subset of Johnston's organ neurons to zones A and B of the brain's antennal mechanosensory centre. It emerges[2,3] that gravity and wind stimuli, which produce somewhat larger, low-frequency or sustained arista motions, are also detected by Johnston's organ neurons, albeit by different subsets of these cells, which project to zones C and E of the antennal mechanosensory centre. (Graphic modified from ref. 13.)

are probably more subtle than those from sound stimuli — and so may activate different subsets of neurons projecting to zones C and E. For example, gravity sensing may require the most sensitive neurons. Alternatively, wind and gravity may segregate perceptually through association with distinctive visual and proprioceptive cues arising from, say, upward versus forward locomotion.

Compartmentalization of Johnston's organ neurons into groups with distinct mechanical sensitivities is reminiscent of the vertebrate inner ear. In the ear, hearing receptor cells also show clear adaptations for fast responses relative to gravity-sensitive receptor cells[9,10]. This similarity extends the list of recently discovered parallels between fruitfly and vertebrate ears, including related genes for specifying ear development[11], the involvement of particular classes of ion channels[3], and evidence for mechanisms that amplify the mechanical input to the sensor[7,12]. Such similarities excite the speculation that hearing organs in fruitflies and vertebrates arose from a sensory structure present in a common ancestor[11], rather than independently as was long thought. ■

Ruth Anne Eatock is at the Eaton-Peabody Laboratory, Department of Otology and Laryngology, Harvard Medical School and Massachusetts Eye and Ear Infirmary, Boston, Massachusetts 02114, USA.
e-mail: eatock@meei.harvard.edu

1. Johnston, C. Q. *J. Microsc. Sci.* **3,** 97–102 (1855).
2. Kamikouchi, A. *et al. Nature* **458,** 165–171 (2009).
3. Yorozu, S. *et al. Nature* **458,** 201–205 (2009).
4. Todi, S. V., Sharma, Y. & Eberl, D. F. *Microsc. Res. Tech.* **63,** 388–399 (2004).
5. Kernan, M. *Pflügers Arch.* **454,** 703–720 (2007).
6. Baker, D. A., Beckingham, K. M. & Armstrong, J. D. *J. Comp. Neurol.* **501,** 756–764 (2007).
7. Göpfert, M. C. & Robert, D. *Proc. R. Soc. Lond. B* **268,** 333–339 (2001).
8. Budick, S. A., Reiser, M. B. & Dickinson, M. H. *J. Exp. Biol.* **210,** 4092–4103 (2007).
9. Ricci, A. J. *et al. J. Neurosci.* **25,** 7831–7839 (2005).
10. Vollrath, M. A. & Eatock, R. A. *J. Neurophysiol.* **90,** 2676–2689 (2003).
11. Boekhoff-Falk, G. *Dev. Dyn.* **232,** 550–558 (2005).
12. Hudspeth, A. J. *Neuron* **59,** 530–545 (2008).
13. Christensen, A. P. & Corey, D. P. *Nature Rev. Neurosci.* **8,** 510–521 (2007).

QUANTUM OPTOELECTRONICS

Swift switch of the strong

Claire Gmachl

How fast can light and matter be made to interact? 'Almost instantaneously' is the answer provided in the latest study of semiconductor structures embedded in an optical microcavity.

In recent decades, few things have changed as rapidly as the way in which we communicate with one another over distance. Starting from basic telephony, we have moved on to the Internet and web-based communication: e-mail, instant messaging, blogging, Facebook and Twitter. Rarely do we realize, however, that we owe the speed and pervasiveness of web-based communication to our ability to switch light on and off swiftly and conveniently in semiconductor devices. Fast-forward into the future, when transferring increasing amounts of data will require manipulation of light at even greater speeds, and at least one crucial question will need to be addressed: how fast can light–matter interactions, the building blocks of switches and light modulators, be turned on and off? On page 178 of this issue, Günter *et al.*[1] show that ultrastrong coupling of light and matter can be turned on practically instantaneously, significantly faster than a single oscillation cycle of the light involved in the process.

In their experiment, Günter and colleagues studied a semiconductor structure comprising 50 quantum wells — thin layers of gallium arsenide (GaAs) — separated by aluminium gallium arsenide (AlGaAs) barriers. This stack of quantum wells was sandwiched between two wider layers of bulk semiconductors (see Fig. 1 of the paper[1] on page 179). Each quantum well confines electrons to two energy sub-bands within the quantum well's conduction band (the uppermost energy band that is separated from the valence band by a forbidden band gap). This two-state system forms the matter part of the light–matter system; the wider cladding layers form an optical microcavity that confines the system's light component, the cavity mode.

An electron in the lower-energy sub-band will be excited into the upper sub-band by absorbing light at the appropriate frequency, whereas an electron in the upper sub-band can revert to the lower sub-band by spontaneously emitting light. The energy difference between the two sub-bands is determined by the thickness of the quantum wells, and was chosen to be 113 millielectronvolts (meV); the period of the corresponding light (the cavity mode) necessary to trigger the intersub-band transition is 37 femtoseconds (1 femtosecond is 10^{-15} seconds). In semiconductors, intersub-band transitions are distinguished from other optical transitions by their strong coupling to light and their narrow spectral width (in this case, 5 meV). This makes them fairly straightforward to observe as narrow and pronounced dips or peaks in light spectra.

The natural state of the semiconductor structure is an undoped state, that is, no electron occupies either the lower or upper sub-bands, and even ambient temperature is not enough to provide a noticeable population of electrons in either sub-band. So to couple light and matter, electrons need to be injected into the conduction sub-bands. The higher the electron density, the stronger the coupling (the coupling strength follows a square-root dependence on the density). In earlier experiments[2–4], electrons were provided through extrinsic doping or an applied voltage.

Observation of the magnitude of the light–matter coupling is done by measuring the

© 2009 Macmillan Publishers Limited. All rights reserved

NATURE|Vol 458|12 March 2009

reflection of broadband light from the microcavity[5–7]. The reflected light spectrum depends on the strength of the coupling. A weakly coupled system exhibits only one reflection resonance — a pronounced dip in the spectrum (at 113 meV in Günter and colleagues' system[1]). Strong and ultrastrong coupling, meanwhile, result in two reflection resonances that are clearly separated from each other by up to 20% of the original resonance's energy, and that are located on either side of this resonance (see Fig. 3 of the paper[1] on page 180).

But let's restate the question: how fast can strong light–matter coupling be turned on or off? To address this question, the authors[1] used a high-energy light pulse of 12-femtosecond duration that could excite electrons from the valence band into the lowest-energy sub-band within that timescale. They found that the strongly coupled light–matter system switches on almost instantaneously, not at its own intrinsic speed (the oscillation of its light field) but at the speed of the light pulse used to pump the electrons (12 femtoseconds compared with 37 femtoseconds): as soon as the electrons reach the conduction sub-bands, the light–matter coupling is established. What's more, the switch is done non-adiabatically rather than adiabatically — that is, the system's coupled quantum state does not emerge gradually but rather switches on abruptly.

Günter and colleagues' work is excellent news in many ways. First, the observation that light–matter coupling can be tuned in less than one cycle of light means that a number of non-adiabatic, non-thermal-equilibrium light–matter phenomena, which belonged solely to the niche of theory, can now be tested in experiments. For example, the observation of the emission of virtual photons lodged in every optical cavity[8] should now be possible. Second, the fact that the experiments are conducted at ambient temperature and pressure makes them accessible to a wider group of experimentalists. Finally, the authors' demonstration of ultrafast switching of ultrastrong light–matter coupling in semiconductor-based structures may, in the long run, prove very useful for communication applications. ■

Claire Gmachl is in the Department of Electrical Engineering, and the Engineering Research Center for Mid-InfraRed Technologies for Health and the Environment, Princeton University, Princeton, New Jersey 08544, USA.
e-mail: cgmachl@princeton.edu

1. Günter, G. *et al. Nature* **458,** 178–181 (2009).
2. Anappara, A. A. *et al.* Preprint at http://arxiv.org/abs/0808.3720 (2008).
3. Anappara, A. A., Tredicucci, A., Beltram, F., Biasiol, G. & Sorba, L. *Appl. Phys. Lett.* **89,** 171109 (2006).
4. Dupont, E., Gupta, J. A. & Liu, H. C. *Phys. Rev. B* **75,** 205325 (2007).
5. Kübler, C. *et al. Appl. Phys. Lett.* **85,** 3360–3362 (2004).
6. Kröll, J. *et al. Nature* **449,** 698–701 (2007).
7. Gaal, P. *et al. Nature* **450,** 1210–1213 (2007).
8. De Liberato, S., Ciuti, C. & Carusotto, I. *Phys. Rev. Lett.* **98,** 103602 (2007).

CONDENSED-MATTER PHYSICS

Pressure for change in metals

N. W. Ashcroft

When is a metal not a metal? When it is under high pressure, if it's lithium or sodium. The strange behaviour of dense forms of these elements exposes difficulties with commonly used models of electronic structure.

The physical laws that govern the behaviour of individual particles of matter can be relatively simple compared with the difficulties of understanding systems that contain large numbers of them. Yet such difficulties must be overcome if one is to understand the subtle quantum physics of the condensed states of matter. So-called simple metals, such as lithium and sodium, have had a pivotal role in studies of metallic solids, because their electronic behaviour is plausibly described using an approximation known as the nearly-free electron model. But two reports[1,2] in this issue reveal that lithium and sodium do something quite remarkable under pressure that can't readily be explained by this approximation — they adopt semiconductor-like forms.

Some 75 years ago, two classic papers[3,4] were published describing the "constitution of metallic sodium". These papers revealed that, under ordinary conditions of temperature and pressure, the electronic structure of sodium is quintessentially free-electron-like. In other words, the behaviour of its valence electrons (the outermost electrons responsible for carrying an electric current) is similar to that of a dense gas of otherwise free electrons. The only differences in behaviour are relatively minor perturbations caused by weak interactions of the valence electrons with the positively charged core structures of their atoms (that is, the ions formed when the valence electrons leave the atoms). Nearly-free electron behaviour was soon recognized as being characteristic of many metals, but especially of lighter elements classified as simple metals. The 'simplicity' of these metals also refers to the spatial arrangements of the crystal lattices formed by the atoms. For example, under ordinary conditions, lithium and sodium often adopt one of the most straightforward of all crystal structures — the body-centred cubic arrangement, consisting of subunits in which eight atoms form the corners of a cube while another sits at the centre of the cube.

But it is increasingly evident that the electronic structures of simple metals under high pressures — and therefore at higher densities of the metals — can be a far cry from those observed under ambient conditions. In the case of lithium, Matsuoka and Shimizu[1] (page 186) demonstrate that, at pressures of about 80 gigapascals (just a little under a quarter that at the centre of the Earth, or roughly 800,000 times atmospheric pressure), the electrical resistance of the element increases abruptly to more than 10,000 times its value at atmospheric pressure. The temperature dependence of the resulting material is then much more characteristic of a semiconductor than of a metal.

Meanwhile, Ma *et al.*[2] (page 182) report that thin slices of sodium (about 3–5 micrometres thick) become transparent to visible light at pressures of about 200 gigapascals (just over half the pressure at the centre of the Earth) in a high-pressure cell. This suggests that the metal has become an insulator, or at least a semiconductor. The authors also obtained spectroscopic data from their dense sodium, which can provide clues on what transformations might have occurred in the crystal structure. These results[2] and those of Matsuoka and Shimizu[1] are surprising, because they cannot be easily explained by historical assumptions about how the crystal structures of simple metals might change under pressure. In particular, there are no special reasons to expect dramatic changes in the conductivities of these metals with increasing pressure, even if the crystal structure undergoes moderate changes as a result of this treatment.

So might these findings[1,2] presage a shift in viewpoint about the electronic structures of high-density metals? Up to now, it has been generally accepted that electrons form patterns of varying electron density around the core structures of their atoms, in response to the electric fields generated by these core structures. Indeed, this long-standing notion — which results in atoms having almost spherical shapes — evolved into the concept of electron shells and subshells in atoms.

But what happens if the average separation of atoms in a solid is significantly reduced (say, by 40%) from the value at ordinary pressure? Under these circumstances, each atom suffers invasion of its territory by its neighbours, which can affect the shapes of the local volumes available to electrons. Furthermore, the core structures of the atoms are forced to overlap; valence electrons, which are already largely excluded from the ion cores, are thus separated even further. To accommodate this increasing exclusion, the valence electrons could instead occupy interstitial locations between the cores, and the crystal structure of the metal might well have to change to accommodate such electron-occupied interstices[5]. Recent extraordinary revelations[6] about the complexity of the structure of sodium at high densities arguably provide ample testimony to the likelihood of such pressure-induced changes.

© 2009 Macmillan Publishers Limited. All rights reserved

In the current reports[1,2], the crystal structures at the onsets of the non-metallic states of lithium and sodium have not been definitively established. Ma *et al.*[2], however, used extensive quantum-mechanical calculations and structural-refinement techniques to predict that sodium adopts an unusual "distorted double-hexagonal close-packed" structure (see Fig. 4a on page 184) at high pressures. This structure is in fact shown to be an insulator at low temperatures. Nevertheless, one cannot yet say with certainty that dense sodium takes this form, because the calculations predict structures at a temperature of absolute zero, whereas Ma and colleagues' measurements were made at room temperature. The calculations also assume that the core structures are infinitely massive. In fact, both lithium and sodium are light elements, and so quantum effects of the atoms associated with their lattice dynamics might well be important for determining the actual structures adopted by the dense metals.

The results[1,2] are not entirely unheralded, as significant changes in the optical characteristics and conductivity of sodium[7] and lithium[8] have previously been noted at pressures not too far from those now described. Nor are the results without some theoretical precedent: about ten years ago it was predicted[9] that lithium might well depart from its simple crystal and electronic structures at higher densities. And predictions for sodium soon followed suit[10]. But what the present results most assuredly demonstrate is the importance of pressure in revealing the limitations of previously hallowed models of solids and their associated electronic behaviour. Furthermore, it is clear that the structures adopted by solids (and even liquids) under pressure will provide stringent tests[11] of our ability to predict both the forms and thermodynamic behaviour of condensed states of matter. ■

N. W. Ashcroft is at the Laboratory of Atomic and Solid-State Physics, Clark Hall, Cornell University, Ithaca, New York 14853-2501, USA.
e-mail: nwa@ccmr.cornell.edu

1. Matsuoka, T. & Shimizu, K. *Nature* **458,** 186–189 (2009).
2. Ma, Y. *et al. Nature* **458,** 182–185 (2009).
3. Wigner, E. & Seitz, F. *Phys. Rev.* **43,** 804–810 (1933).
4. Wigner, E. & Seitz, F. *Phys. Rev.* **46,** 509–524 (1934).
5. Rousseau, B. & Ashcroft, N. W. *Phys. Rev. Lett.* **101,** 046407 (2008).
6. McMahon, M. I. *et al. Proc. Natl Acad. Sci. USA* **104,** 17297–17299 (2007).
7. Syassen, K. in *Proceedings of the International School of Physics "Enrico Fermi": High Pressure Phenomena* Vol. 147 (eds Chiarotti, G. L., Hemley, R. J., Bernasconi, M. & Ulivi, L.) 266–268 (IOS Press, 2002).
8. Deemyad, S. & Schilling, J. S. *Phys. Rev. Lett.* **91,** 167001 (2003).
9. Neaton, J. B. & Ashcroft, N. W. *Nature* **400,** 141–144 (1999).
10. Christensen, N. E. & Novikov, D. L. *Solid State Commun.* **119,** 477–490 (2001).
11. Hemley, R. J. & Ashcroft, N. W. *Physics Today* No. 8, 26–32 (1998).

CELL BIOLOGY

Detached membrane bending

Hélène Barelli and Bruno Antonny

Cells use various protein complexes to remodel membrane-bound organelles. *In vitro* reconstitution of the activity of one such complex, ESCRT-III, shows that it promotes membrane bending in an unconventional way.

Many transmembrane proteins — including those that span the cell membrane — are degraded within cellular organelles called lysosomes once their work is done. Reaching the lysosomes is not simple, however, and involves several membrane-vesicle trafficking steps[1]. On page 172 of this issue, Wollert *et al.*[2] elucidate the role of the components of the ESCRT-III protein complex in one stage of this process. The authors' results are of interest not just for the insights they provide into the processes of intracellular trafficking and protein degradation, but also because ESCRT proteins mediate the budding of certain viruses, including HIV-1, as well as the separation of the two daughter cells at the end of cell division.

So what is the itinerary for the journey of a cell-membrane protein to the lysosomes? First, small portions of cell membrane invaginate and become detached to form organelles called endosomes, incorporating the proteins. The endosome further invaginates internally into its own lumen to form several intraluminal vesicles (ILVs). Endosomes with their ILVs are called multivesicular bodies, and eventually fuse with lysosomes, where ILVs and their contents are degraded by the cocktail of lysosomal enzymes.

Compared with other membrane-deformation events that occur in the cell, ILV formation is puzzling (Fig. 1). Vesicles generally bud from organelles (such as the Golgi complex) in the opposite direction to ILV formation: a protein coat assembles on the cytoplasmic side of the organelle membrane to form an outer shell and impose deformation towards the cytoplasm[3]. Once the vesicle has become detached from the membrane, the protein coat disassembles and is reused for another round of budding.

Given the opposite topology of ILVs, budding away from the cytoplasm, could it be that ESCRT-III sculpts the membrane from inside the bud? After all, induction of curvature from the inside has precedents: some proteins with

© 2009 Macmillan Publishers Limited. All rights reserved



50 YEARS AGO

A Congress was held in Singapore during December 2–9 to celebrate "the Centenary of the formulation of the theory of Evolution by Charles Darwin and Alfred Russel Wallace and the Bicentenary of the publication of the tenth edition of the 'Systema Naturae' by Linnaeus". It was particularly fitting that this Congress should have been held in Singapore for ... it directed special attention to the work of Wallace, who was one of the greatest biologists ever to have worked in south-east Asia ... Prof. Haldane then delivered his presidential address ... The president emphasised the stimuli gained by Linnaeus, Darwin and Wallace through working in peripheral areas where lack of knowledge was a challenge. He suggested that the next major biological advance may well come for similar reasons from peripheral places such as Singapore, or Calcutta, where this challenge still remains and where the lack of complex scientific apparatus drives biologists into different and long-neglected fields of research.
From *Nature* 14 March 1959.

100 YEARS AGO

On Monday evening Dr. M. A. Stein read before the Royal Geographical Society a paper on his geographical and archaeological explorations in Chinese Turkestan in 1906–8 ... He was greatly desirous of examining a secret store of ancient manuscripts which had been accidentally discovered by a Taoist priest in the Caves of the Thousand Buddhas ... These were piled up without any sort of order to a height of 10 feet, and comprised not only written documents, but fine paintings on silk and cotton, ex-votos in all kinds of silk and brocade, and streamers in various fabrics. Dated documents showed that the chamber must have been walled up about 1000 A.D., but some of the records dated back so far as the third century A.D.
From *Nature* 11 March 1909.

50 & 100 YEARS AGO

convex shapes adhere to lipid membranes to form protrusions[4], and some toxins, such as Shiga toxin, invade the cell by inducing their own uptake into cell-membrane invaginations[5] (Fig. 1). But the problem with ESCRT-III pertains to its recycling; if they act from within the bud, ESCRT-III proteins would be trapped in the ILV and so would be used only once.

Wollert *et al.*[2] asked whether ESCRT-III might act outside the invagination, which would allow its subsequent recycling. To address this question, they used the well-established approach of biochemical reconstitution. In short, they added purified recombinant yeast ESCRT-III proteins to artificial, giant unilamellar (single-membrane) vesicles (GUVs) that resemble cellular membranes, and studied them using light microscopy. Their observations are most informative.

On incubation with three components of ESCRT-III (Vps20, Snf7 and Vps24), the GUVs become filled with internal vesicles that are clearly detached from the surrounding membrane. This finding suggests that these three proteins are sufficient to promote the entire process of ILV formation — from the induction of membrane curvature to the scission of the bud neck.

In a second set of experiments, the authors added two further ESCRT-III proteins to the mixture: Vps4, an ATPase enzyme that mediates the disassembly of ESCRT-III proteins, and Vps2, the final component of ESCRT-III, which connects Vps4 to the rest of the ESCRT-III proteins. This time, Wollert *et al.* found that a second wave of ILVs followed the first, being identified by having their contents stained with a green dye. So, although Vps4 is not involved in the formation of ILVs, its ATPase activity seems to allow the recycling of the ESCRT-III machinery.

Wollert and colleagues' strategy seems so straightforward that one wonders why it was not used before. But biochemical reconstitution with five pure proteins and a model lipid membrane is more difficult than it sounds. To prevent self-aggregation of each protein, the authors had to use specifically engineered proteins. Moreover, the absolute concentration of the proteins, their relative ratios and the lipid composition of the GUVs were all carefully chosen. Of course, a GUV is not a cellular organelle, but these precautions highlight the relevance of the effects observed.

The results[2] support the idea that, indeed, ESCRT-III drives membrane invagination and fission while itself remaining at the door (or neck) of the forming ILV. But our pen trembles when we consider drawing an explicit model. The structures of some of the proteins involved in membrane deformation speak for themselves — take the triskelion shape of the clathrin protein, or the banana shape of BAR-domain proteins, for example. This is not, however, the case for ESCRT-III proteins, which share a simple elementary fold that can be envisaged as a small, flat jigsaw puzzle, with a basic charge well suited for interaction with the negatively charged membrane lipids[6].

Recent work[7–10] suggests that ESCRT-III proteins assemble in an ordered manner at the surface of lipid membranes, where they form spiral filaments (Fig. 2). The spiral shape allows ESCRT-III to act as a fence encircling cargo proteins en route to the nascent ILV. But what drives membrane invagination?

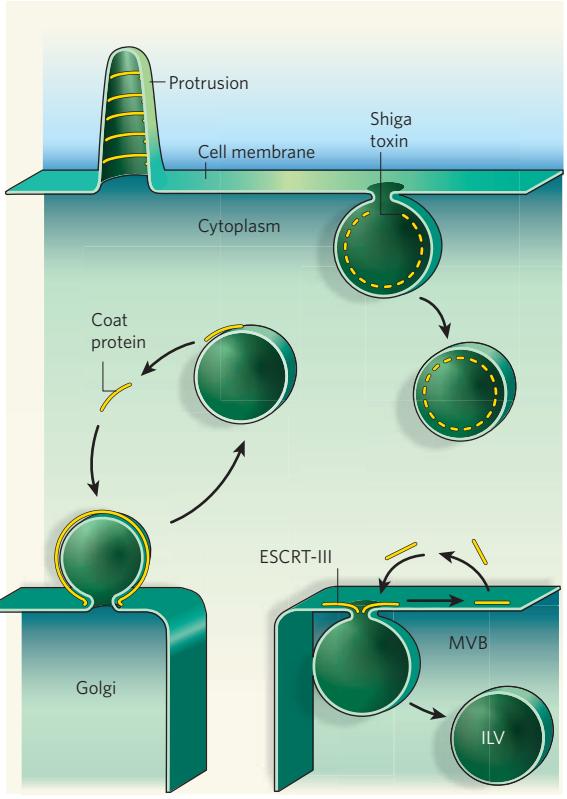


Figure 1 | Membrane-deformation events within the cell. Membrane-bound vesicles can form in different ways. Protein coats form an outer shell around transport vesicles, such as those budding off the Golgi complex, and can be recycled in the cytoplasm. Proteins with a convex shape sculpt protrusions from the inside but can be reused because no membrane fission occurs. Shiga toxin deforms the cell membrane, itself becoming trapped in the invagination. Wollert and colleagues' data[2], based on *in vitro* reconstitution studies, indicate that the ESCRT-III complex might be sufficient to drive formation of intraluminal vesicles (ILVs) within a multivesicular body (MVB) from the neck of the forming bud, without entering it; thus, ESCRT-III proteins can be recycled for another round of ILV formation.

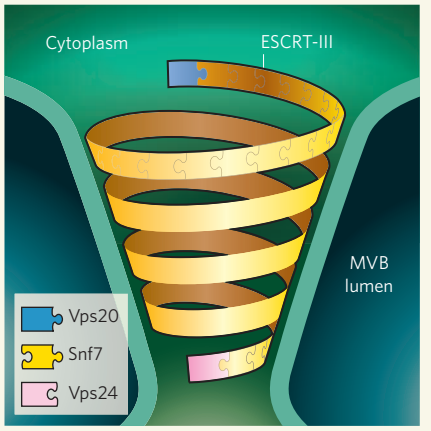


Figure 2 | ESCRT-III and membrane curvature. The ESCRT-III complex consists mainly of three proteins — Vps20, Snf7 and Vps24 — which are thought to form a spiral structure. How the spiral shape of ESCRT-III relates to its ability to promote ILV budding and fission remains a matter of debate.

If the spiral is smooth, and if the basic face of the monomers is parallel to the main plane of each turn (like a flat, coiled piece of rope; see Fig. 6 on page 176), it is hard to imagine how it could drive invagination: the spiral should happily sit on a flat, negatively charged membrane. But if the basic surface is tilted, and the spiral has some stiffness, each turn might come out of the previous one, both to escape mechanical constraint and to interact with the membrane (Fig. 2). Under these circumstances, a conical invagination should form. But the spiral should not invade the neck so deeply as to create a narrow, empty lipid tubule ready to undergo scission. Intriguingly, Vps24 seems both to act as a stop signal for the elongation of the spiral[10] and to favour the ILV fission step[2]. Such speculations highlight the fact that ESCRT-III is so different from other protein machinery involved in intracellular trafficking that Wollert and colleagues' study[2], and those of others[7–10], are probably just the beginning of many surprises yet to come. ■

Hélène Barelli and Bruno Antonny are at the Institut de Pharmacologie Moléculaire et Cellulaire, CNRS et Université de Nice, 660 route des Lucioles, 06560 Valbonne, France.
e-mail: antonny@ipmc.cnrs.fr

1. Hurley, J. H. & Emr, S. D. *Annu. Rev. Biophys. Biomol. Struct.* **35,** 277–298 (2006).
2. Wollert, T., Wunder, C., Lippincott-Schwartz, J. & Hurley, J. H. *Nature* **458,** 172–177 (2009).
3. Antonny, B. *Curr. Opin. Cell Biol.* **18,** 386–394 (2006).
4. Mattila, P. K. *et al. J. Cell Biol.* **176,** 953–964 (2007).
5. Romer, W. *et al. Nature* **450,** 670–675 (2007).
6. Muziol, T. *et al. Dev. Cell* **10,** 821–830 (2006).
7. Teis, D., Saksena, S. & Emr, S. D. *Dev. Cell* **15,** 578–589 (2008).
8. Hanson, P. I., Roth, R., Lin, Y. & Heuser, J. E. *J. Cell Biol.* **180,** 389–402 (2008).
9. Lata, S. *et al. Science* **321,** 1354–1357 (2008).
10. Saksena, S., Wahlman, J., Teis, D., Johnson, A. E. & Emr, S. D. *Cell* **136,** 97–109 (2009).

© 2009 Macmillan Publishers Limited. All rights reserved

OBITUARY

Xiangzhong (Jerry) Yang (1959–2009)

Leading light in animal cloning.

Jerry Yang, who died on 5 February in Boston, Massachusetts, made exceptional contributions to research on animal biotechnology and cloning, and was a prominent figure in the scientific dialogue between the United States and China. He was only 49 when he died, finally succumbing to cancer of the salivary gland. Losing battles, however, was not Yang's way of life.

He was born in 1959, in China, and barely survived the famine of 1959–60. His parents were poor pig farmers in the tiny village of Dongcun, about 500 kilometres south of Beijing, but in the late 1970s the award of a place at Beijing Agricultural College set Yang on the road to a wider world. A scholarship to Cornell University in New York followed, where he took a master's degree and completed his PhD in reproductive physiology. Success in his subsequent research at Cornell led to his being hired by the University of Connecticut in 1996, where he held various positions, including director of the Center for Regenerative Biology. It was no mean achievement to have risen from such humble beginnings to conquer the Chinese education system, and then to have won major grants and promotions in the demanding US tertiary research system.

In another respect, Yang was fortunate in having a well-prepared mind at a momentous juncture in the history of research on animal embryology. The birth of the cloned sheep Dolly was announced in 1997, and had an influence far beyond the scientific community concerned. Her birth was achieved through the procedure of somatic-cell nuclear transfer (SCNT), in which the nucleus — and so the genetic complement — of a mammary-gland cell from a mature ewe was transferred into an unfertilized sheep oocyte or egg cell, which was then allowed to develop before being implanted in a recipient ewe and coming to term. Cloning from an adult cell rarely works; it requires hundreds of nuclear transfers to produce a viable animal. Nor is it successful in humans. But it promised to revolutionize animal breeding, through the cloning of prize livestock. And it had the potential to create a new branch of regenerative medicine, one using the patient's own immune-compatible cells that could be reprogrammed by cloning to produce pluripotent stem cells for the repair of damaged and diseased tissues.

Yang was in the thick of the ensuing period of intensive research. In papers that appeared between 2000 and 2006, his work confirmed that — in cattle, rabbits, pigs, mice and the swamp buffalo — fully differentiated adult



J. SAUER/AP PHOTO

somatic cells can be reprogrammed into developmentally normal embryonic cells capable of forming fully viable offspring. Moreover, he showed that first- and second-generation clones have normal fertility, which was particularly informative for sustained breeding of genetically valuable cloned bulls.

Also in early work, Yang settled the debate on telomere shortening — the serial reduction in length of the ends of chromosomes as cells divide — which had been proposed as a limiting factor in the viability of cloned animals. Telomere length was unaffected by nuclear-transfer procedures, and no shortening was found in offspring cloned from even aged adults. Yang's work on cattle and many other species catapulted the United States into the forefront of animal cloning, with many biotechnology companies examining the commercial opportunities for investment in animal breeding, and for protein production in transgenic and cloned animals.

In his basic research, Yang explored the cell biology of nuclear reprogramming. Notably — and contrary to expectations at the time — in 2006 he controversially reported that cloning efficiency increased with increasing differentiation within the hierarchy of progenitor cells in the haematopoetic (blood-cell) system; put technically, terminally differentiated post-mitotic granulocytes yielded cloned pups with greater efficiency than did the earlier adult stem-cell or progenitor types. Furthermore, he used molecular microarray analyses to demonstrate that adult stem cells could be reprogrammed by SCNT to acquire an embryonic-like profile of expressed genes, and he showed that the gene profile was closer to that observed in embryos fertilized *in vivo* than to that in embryos fertilized *in vitro*. But the SCNT embryos had markedly reduced developmental capacity. The apparent disconnect between embryonic gene expression and embryo viability was a particular preoccupation towards the end of his life.

In 2006, Yang's life's work, and indeed a whole field of research, was put in jeopardy by the exposure of fraud by the South Korean cloning team led by Woo Suk Hwang. Hwang falsely claimed to have achieved the highly efficient production of numerous human SCNT embryonic stem-cell lines. Yang was a strong advocate for human SCNT, given its promise in 'therapeutic cloning' by providing immune-compatible tissues for cell replacement, and he was deeply concerned that this episode would seriously damage the field of regenerative medicine. In high-profile articles, he argued the case both for continued research in cloning by SCNT and for increased vigilance of scientific fraud.

International collaboration was the other main strand of Yang's professional life, and he was a prominent proponent of global partnerships in science and medicine. He achieved the feat of being respected in both the academic and political sectors in China, and was a persuasive advocate for US–China cooperation. He pursued his goals with typical tenacity, an example being his advocacy for an Asian–Pacific network of stem-cell scientists. His case was simple: Asia is a region where political and religious ideologies are not antagonistic towards human stem-cell science. He proposed the creation of a network of regenerative-medicine institutes in China, with an international oversight committee to guarantee transparency, trust and compliance with established ethical standards, and he worked hard to establish funding and to create a suitable organization for this network. He hoped this model would be a beacon to guide the formation of other international scientific networks.

Jerry Yang had a passion and determination that marks the very best in the traditions we admire and respect. He was an inspiration to the young Chinese and US scientists who studied with him, and who have since populated many different laboratories. He was a terrific person to spend time with: I, like many others, regret that I cannot now shoot the breeze with him on animals, embryos and biotechnology.

Alan Trounson
Alan Trounson is at the California Institute for Regenerative Medicine, 210 King Street, San Francisco, California 94107, USA.
e-mail: atrounson@cirm.ca.gov

© 2009 Macmillan Publishers Limited. All rights reserved

Q&A

EVOLUTIONARY BIOLOGY

Speciation

Andrew P. Hendry



Darwin200

On the Origin of Species ... the title of Charles Darwin's great work of 1859 seemed to promise a solution to this "mystery of mysteries". Although we now know vastly more about speciation than we did 150 years ago, the one mystery has become many — and the possible solutions have multiplied.

Are species real?

The entities we call species — definable biological groups of distinct lineage and with potentially independent futures — are certainly real. But the term 'species' as a level of biological classification is more ambiguous and amorphous; for example, the species identified by criteria typically applied to one branch of life (or by one set of biologists) would not always be similarly identified by criteria typically applied to other branches of life (or by another set of biologists). In this sense, the classification level 'species' is partly a construct developed for the convenience of biologists in organizing and simplifying the natural world. Darwin himself, in *On the Origin of Species*, was of this opinion: "In short, we shall have to treat species in the same manner as those naturalists treat genera, who admit that genera are merely artificial combinations made for convenience. This may not be a cheering prospect; but we shall at least be free from the vain search for the undiscovered and undiscoverable essence of the term species." Although Darwin has been criticized for this ambivalence, modern research has, in many respects, returned us to his original interpretation.

What taxonomic groups are we talking about?

All of them. The diversity of life around us is the product of evolution that has generated different life forms embarking on different evolutionary trajectories. This statement applies all the way from the big (polar bears versus grizzly bears) to the very small (different strains of bacteria or viruses), and across the entire tree of life. The hitch is that difficulties arise in deciding how to define species within and across these vastly different groups.

How are species delineated by biologists?

A universal 'species concept' has long been sought. The hope has been for a basic set of rules that can be broadly applied across many branches of life so as to clearly, objectively and comparably delineate separate species. Proposing species concepts has proven appealing, and has become almost a cottage industry — more than two dozen have been developed. Although none is universally accepted, one of the most popular is the 'biological species concept'.



THE NATURAL HISTORY MUSEUM, LONDON

Figure 1 | Beetle mania: one estimate is that some 350,000 species have been described.

What is the 'biological species concept' (BSC)?

The BSC holds that species are groups of actually or potentially interbreeding individuals that are reproductively isolated from other such groups (that is, they exchange few genes). The BSC is sometimes interpreted to imply the extreme situation where two groups are separate species only when successful hybrids cannot ever be produced — and any two such groups certainly are separate species. But many other groups that are widely accepted to represent separate species frequently violate this strict criterion; for example, some estimates hold that 25% of all plant species and 10% of all animal species hybridize successfully with at least one other species. Partly for this reason, the BSC is often relaxed to the point that different groups are considered separate species if they can maintain their genetic integrity in nature. This more useful, albeit more ambiguous, criterion allows for some genetic exchange (gene flow) between species as long as they do not become homogenized.

What are the challenges to the BSC (and the alternatives)?

One is that it is very difficult to apply the BSC to groups that are geographically separated ('allopatric'), because it is hard to know if these groups would collapse into a single species if they did somehow come into contact (become 'sympatric'). This ambiguity is one reason for the qualifier 'potentially' in the BSC definition given above. For example, squirrels in New York's Central Park do not interbreed with those in nearby New Jersey, but perhaps only because of the impassable expanse of water and concrete between them. A common route around this problem is to bring allopatric groups into artificial contact, such as in the laboratory. Ambiguity still persists here, however, because the observed mating patterns may not reflect those that would occur in nature. A second challenge is that the BSC cannot be applied to *most* organisms, including those without sexual reproduction, such as viruses and some microbes, and those that are now extinct. For these reasons, and others, an alternative set of species concepts revolves around the magnitude of morphological or genetic differences between groups. The difficulty for

© 2009 Macmillan Publishers Limited. All rights reserved

these concepts is in deciding just how big a difference is sufficient to elevate two groups to the status of separate species.

At what point do groups become species?

Speciation is best thought of not as a specific endpoint (the above answers highlight the difficulty of establishing an appropriate threshold), but rather as an accumulation of reproductive isolation and of morphological/genetic differences through time. This emphasis on speciation as a *process* has helped to refocus research towards the factors that promote and constrain 'progress' towards (or from) speciation. This focus also avoids the philosophical awkwardness that the consideration of speciation as an endpoint requires that offspring are one day born who are not the same species as their parents. (Although I will describe below how this can sometimes happen.)

How many species are out there?

Approximately 1.5 million species are currently recognized by taxonomists to the point of having been awarded a Linnean binomial, such as us — *Homo sapiens*. Associated with each official species description is a 'type' specimen, considered representative of the species (curiously, humans have never had a designated type specimen, despite attempts by American palaeontologist Edward Drinker Cope to have himself so designated). One and a half million sounds like a lot, but many other species have yet to be described, or even discovered; and so the total that might warrant taxonomic designation could be upwards of 10 million. And, of course, vastly more species have gone extinct than exist today. Within animals, the most speciose group is insects. And within insects, the most speciose groups are those that feed on plants. It's said that the evolutionary biologist J. B. S. Haldane, on being asked by theologians what he had learned about the Creator, stated that He must have had "an inordinate fondness for beetles" (Fig. 1).

How quickly do species arise?

Starting with Darwin, and continuing until very recently, speciation has been considered to be usually a very slow process — perhaps requiring millions of years. But this cannot always be true given, for example, the many hundreds of cichlid fish species in Lake Victoria, Africa, that apparently arose in less than 15,000 years. On a similar time frame, numerous fish and bird species have arisen in the Northern Hemisphere since the end of the last glaciation. It has also been argued — although not without controversy — that noteworthy levels of reproductive isolation can begin to accumulate within only a few dozen generations; however, the attainment of unassailable and irreversible reproductive barriers will almost always take much longer. An exception is the almost instantaneous speciation that occurs when genomes multiply (polyploidy) and individuals with different ploidy levels cannot successfully interbreed. An example is the common cordgrass plant *Spartina anglica* in Britain, a product of chromosome doubling associated with hybridization between native (*S. maritima*) and introduced (*S. alterniflora*) species.

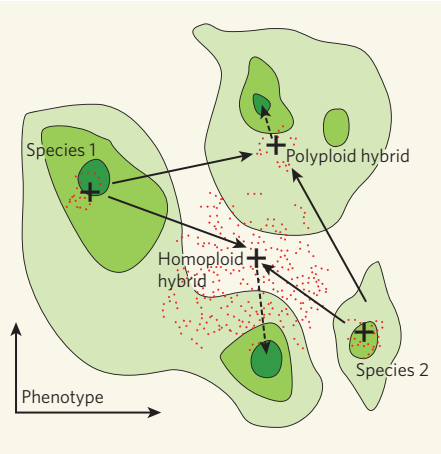


Figure 2 | Hybrid speciation. This example of hybrid speciation, depicted on an 'adaptive landscape', shows evolutionary fitness (contours, with darker green representing higher fitness) as a function of two different phenotypic traits (horizontal and vertical dimensions). Several high-fitness peaks are possible, but initially only two are occupied: by species 1 and species 2, with their average values shown as crosses and the range of individuals as dots. Species 1 and 2 then hybridize, either with (polyploid) or without (homoploid) changes in chromosome number, and the resulting combinations of traits can come into the range of attraction of a new adaptive peak. Adaptation of the hybrids to that peak can then generate a new species. (Modified from J. Mallet *Nature* 446, 279–283; 2007.)

What is the engine of speciation?

For most of the twentieth century, the dominant view of speciation was that random genetic differences accrue gradually between allopatric populations, eventually causing genetic incompatibilities that prevent successful interbreeding. More recently has come the realization that most of these differences evolve by selection. Moreover, selection can cause reproductive barriers other than strict genetic incompatibilities, and this can occur in sympatry or allopatry (more about this later). Of particular recent interest is how adaptation to different environments drives reproductive isolation, a process now called 'ecological speciation'.

... and ecological speciation works how?

One of Darwin's insights, supported by countless subsequent studies, is that populations in different environments show adaptive differences in traits that thereby improve survival and reproductive success. In some cases, this adaptive divergence can lead to reproductive isolation. For example, traits undergoing adaptive divergence might influence mating preferences, such that individuals adapted to different environments now reject each other as mates. Moreover, any hybrid offspring that are produced might die because they aren't well adapted to either environment. Although Darwin wasn't explicit about such a process, it is clear that he too considered adaptive divergence to be the engine of speciation. At present, ecological speciation is thought to have been a major player in the diversity of life. Some well-known examples for vertebrates include Darwin's finches of the Galapagos Islands, threespine stickleback fishes in the Northern Hemisphere and the hyperdiverse cichlids of Africa.

Does speciation require geographical isolation (allopatry)?

Yes. Wait, I mean no. Or perhaps maybe. Allopatric versus sympatric speciation remains one of the greatest controversies in evolutionary biology. Until recently, it was widely considered almost impossible for speciation to proceed entirely in sympatry. (Definitions of sympatry vary widely, but generally accepted contexts would include insects in a field or small forest, fishes in a small lake, or birds or plants on a small island.) And any suggestions to the contrary were quashed by the reigning authorities, particularly Ernst Mayr. Over the past ten years, however, a flurry of mathematical models has demonstrated that ecological speciation is possible in sympatry, and several putative examples have emerged. These include insects adapted to different host plants, fishes adapted to different lake habitats, birds adapted to different food types, and plants adapted to different soil types or pollinators. Most biologists, however, continue to argue that sympatric speciation is relatively rare.

Why is the anti-sympatry sentiment so strong?

Part of the reason may be different standards of proof. Allopatric speciation is usually taken to be the parsimonious default — an appeal to the simplicity of Occam's Razor. Sympatric speciation is thus accepted only when allopatric speciation can be completely ruled out. But this is virtually impossible to do, given that the distributions of organisms at present will rarely reflect their distribution at the time of speciation. And, of course, nature is probably not parsimonious, thus dulling Occam's Razor.

Can competition contribute to sympatric speciation?

Yes, at least under some conditions. Competition for shared resources in sympatry can favour a population splitting into two groups so as to reduce competition, a process some authors have called 'adaptive speciation' or 'competitive speciation'. This process is certainly known to accelerate divergence when formerly allopatric groups come into secondary contact. What remains uncertain is the extent to which competition might

© 2009 Macmillan Publishers Limited. All rights reserved

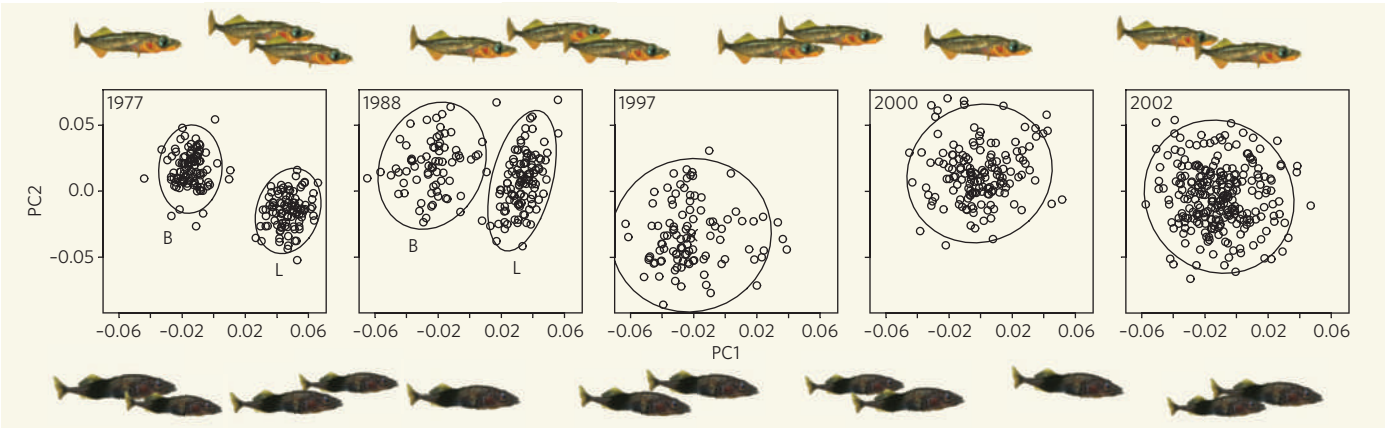


Figure 3 | Reverse speciation. Threespine stickleback fishes in Enos Lake, British Columbia, Canada, historically existed as two distinct groups, one adapted to bottom feeding (benthic, B) and the other to open-water feeding (limnetic, L). The figure shows individuals (small open circles) of each group in morphological space as defined by two orthogonal axes (PC1 and PC2). Historically, very few intermediate (hybrid) individuals were present, despite the two forms — shown here above (L) and below (B) the plots — occurring side-by-side in the same small lake. More recently, however, the two forms have merged into a single group that shows a high degree of morphological and genetic variation. This species fusion may have been the result of human influences. (Modified from E. B. Taylor *et al. Mol. Ecol.* 15, 343–355; 2006. Fish images courtesy E. Cooper, WWF.)

drive sympatric speciation in the absence of an earlier period of allopatry.

... and can hybridization likewise contribute?

Yes. As described above, speciation is usually considered to be the accumulation of morphological/genetic differences and of reproductive isolation — that is, the fission of a single ancestral group into multiple descendant groups. But new species can also arise through fusion, such that hybrids between existing species sometimes embark on separate evolutionary trajectories from their ancestors (Fig. 2). This 'hybrid speciation' is certainly important in plants, where more than 10% of species, perhaps many more, may have a hybrid origin. Hybridization is also increasingly implicated in the origin of some animal species. Many cases of hybrid speciation involve differences in ploidy (as mentioned above), but others do not. For example, researchers have been able to recreate many features of existing hybrid species of sunflowers and butterflies through artificial crosses between existing species, without any differences in ploidy.

Are there 'speciation genes'?

As with most of biology, the study of speciation has gone genomic. One question centres on the number and type of genes that cause reproductive isolation. Some studies have obligingly found genes that differ between species and cause strong reproductive barriers. Speciation may here involve very few genes, although it is often unclear which differences drove speciation and which simply accumulated after the fact. Ecological speciation, in contrast, proposes that reproductive isolation is the result of adaptive divergence, a process that is often caused by many genes. In these cases, 'speciation genes' may be so numerous as to complicate their accurate detection and interpretation. Studies of the genetic basis of reproductive barriers are nonetheless greatly enriching our understanding of speciation.

Can humans promote speciation?

Yes. Humans sometimes sunder the range of a formerly continuous population, which may then promote independent evolutionary trajectories. Humans can also provide new and distinct habitats that initiate ecological speciation. One example is the introduction of plants to sites where they are not native, followed by the evolution of new insect 'host races' that specialize on those plants. These host races can show considerable, although not complete, reproductive isolation from each other in less than 100 years. An example is the recent origin of a host race of tephritid fruitfly (*Rhagoletis*) on brushy honeysuckle (*Lonicera* spp.), the latter having been introduced from Asia to North America. This same fly also answers the above questions regarding how quickly speciation can commence (within a few decades) and how hybridization can form new species (the new host race was initially a hybrid between two other *Rhagoletis* species).

Can humans constrain speciation?

Again, yes. Humans can bring together partially divergent groups that haven't yet accumulated enough differences to prevent their fusion. Humans can also hamper ecological speciation by reducing the distinctiveness of formerly separate resources (habitats or food types), such as by introducing an intermediate and overlapping resource. And human activities can degrade the transmission of mating signals that keep species separate, for instance by reducing water clarity for fishes that choose mates based on colour. So far, human influences have been implicated in 'reverse speciation' or 'de-speciation' in each of these contexts, one example being stickleback fishes in British Columbia, Canada (Fig. 3).

What's next?

Sympatric speciation remains one of the most conspicuous current battlegrounds, and that debate might not wrap up soon. For me, however, the most refreshing voices are those arguing for less concern over the geography of speciation and for more emphasis on its specific drivers. Here, future work will increasingly examine the role of natural selection, including that driven by competition, and also sexual selection (selection to improve the chances of mating). Sexual selection clearly contributes to speciation, but the question is how often it does so without the collaboration of natural selection (or vice versa). Finally, genome studies are sure to increase our understanding of how natural selection acts on regions of the genome that contribute to speciation. In speciation, as in other fields of study, it seems that the more we know, the more we know we don't. Thus, although we now know vastly more than Darwin thought he *didn't* know 150 years ago, the remaining uncertainties are no less glaring. I wonder if we will say the same 150 years from now, or if the questions of speciation will finally have been answered — or perhaps supplanted by new interests. ■

Andrew P. Hendry is in the Redpath Museum and the Department of Biology, McGill University, 859 Sherbrooke Street West, Montreal, Quebec H3A 2K6, Canada.
e-mail: andrew.hendry@mcgill.ca

FURTHER READING
Coyne, J. A. & Orr, H. A. *Speciation* (Sinauer, 2004).
Bolnick, D. I. & Fitzpatrick, B. M. Sympatric speciation: Models and empirical evidence. *Annu. Rev. Ecol. Evol. Syst.* **38,** 459–487 (2007).
Mallet, J. Hybrid speciation. *Nature* **446,** 279–283 (2007).
Seehausen, O., Takimoto, G., Roy, D. & Jokela, J. Speciation reversal and biodiversity dynamics with hybridization in changing environments. *Mol. Ecol.* **17,** 30–44 (2008).
Rundle, H. D. & Nosil, P. Ecological speciation. *Ecol. Lett.* **8,** 336–352 (2005).
Rieseberg, L. H. & Wills, J. H. Plant speciation. *Science* **317,** 910–914 (2007).
Hendry, A. P., Nosil, P. & Rieseberg, L. H. The speed of ecological speciation. *Funct. Ecol.* **21,** 455–464 (2007).

© 2009 Macmillan Publishers Limited. All rights reserved

ARTICLES

The neural basis of *Drosophila* gravity-sensing and hearing

Azusa Kamikouchi[1,2,3]*, Hidehiko K. Inagaki[2]*†, Thomas Effertz[1,4], Oliver Hendrich[1,4], André Fiala[4,5], Martin C. Göpfert[1,4] & Kei Ito[2]

The neural substrates that the fruitfly *Drosophila* uses to sense smell, taste and light share marked structural and functional similarities with ours, providing attractive models to dissect sensory stimulus processing. Here we focus on two of the remaining and less understood prime sensory modalities: graviception and hearing. We show that the fly has implemented both sensory modalities into a single system, Johnston's organ, which houses specialized clusters of mechanosensory neurons, each of which monitors specific movements of the antenna. Gravity- and sound-sensitive neurons differ in their response characteristics, and only the latter express the candidate mechanotransducer channel NompC. The two neural subsets also differ in their central projections, feeding into neural pathways that are reminiscent of the vestibular and auditory pathways in our brain. By establishing the *Drosophila* counterparts of these sensory systems, our findings provide the basis for a systematic functional and molecular dissection of how different mechanosensory stimuli are detected and processed.

The fruitfly *Drosophila melanogaster* responds behaviourally to gravity and sound. When tapped down in a vial, the flies tend to walk up against the Earth's gravitational field, a directed behaviour that is known as negative gravitaxis or anti-geotaxis[1–3]. When exposed to male courtship songs, females reduce locomotion whereas males start chasing each other, forming so-called courtship chains[4,5]. Both *Drosophila* gravitaxis and sound communication have long been prime paradigms for the genetic dissection of behaviour[1–5], but the underlying sensory mechanisms are poorly understood. The human ability to sense gravity and sound relies on specialized vestibular and auditory organs in our inner ear[6,7]. In the fly, the ability to hear has been ascribed to the antenna[5,8–14]: the club-shaped third segment and the distal arista (formed by the fourth to sixth segments) of the antenna sympathetically vibrate in response to acoustic stimuli and, analogous to our eardrum, serve the reception of sound[12,14]. Vibrations of this antennal receiver are picked up by Johnston's organ (JO), a chordotonal stretch-receptor organ with ~480 primary sensory neurons in the second segment of the antenna (Fig. 1a). These JO neurons have also been surmised to have a role in gravity sensing[2,15]. The antennal receiver of the fly is predicted to deflect in response to gravitational forces (see Supplementary Information footnote 1), but physiological evidence exploring the role of JO neurons in gravity sensing has not been reported so far.

Here we examine the role of *Drosophila* JO neurons in gravity and sound detection. It has been shown that the JO neurons of the fly can be anatomically categorized into five subgroups, A–E, each of which targets a distinct area of the brain[13]. Whether this anatomical diversity is paralleled by function, however, has remained unclear[16]. We show that JO neuron subgroups are functionally specialized in that they preferentially respond to distinct types of antennal movement. We further show that this functional diversity reflects distinct behavioural requirements, with different JO neuron subgroups being needed for the response of flies to gravity and sound. These neural subgroups differ genetically and feed into distinct neural pathways in the brain. We have traced these newly identified sensory pathways and provide tools to dissect their function.

Monitoring neural activities in JO

To assess directly neural activities in *Drosophila* JO caused by the antennal receiver movement, we have developed a live fly preparation that affords access to intracellular calcium signals in JO neurons through the cuticle of the antenna (Fig. 1a, b). An intact fly was mounted under a coverslip with the first and second antennal segments immobilized to prevent muscle-based antennal movements. The antennal receiver was kept freely moving, as was confirmed by laser Doppler vibrometric measurements of their mechanical fluctuations[17]. We mechanically actuated the antennal receiver by means of electrostatic force[17–19] (Fig. 1a and Supplementary Fig. 1a), and expressed a genetically encoded calcium sensor in JO neurons via the yeast-derived *GAL4/UAS* gene expression induction system, in which expression of reporter genes fused under *UAS* is activated specifically in the cells that express Gal4 (ref. 20). To distinguish mechanically evoked calcium signals from possible movement artefacts, we used the sensor cameleon 2.1 (Cam2.1)[21,22], which allows for ratiometric measurements of calcium-induced fluorescence resonance transfer (FRET) between enhanced cyan fluorescent protein (eCFP) and enhanced yellow fluorescent protein (eYFP).

When we expressed *cam2.1* in essentially all JO neurons by means of the *F-GAL4* driver[9] (JO-all > *cam2.1*), antennal movement evoked reciprocal changes in eCFP and eYFP fluorescence (Fig. 1c). These signals were largely reduced when *cam2.1* was expressed in homozygous *nanchung* (*nan*^{36a}) mutants[9], but not in heterozygous controls (Supplementary Fig. 1b). Like sound-evoked potentials in the antennal nerve of flies[9], mechanically evoked calcium signals in JO neuron somata thus depend on the transient receptor potential vanilloid (TRPV) channel Nanchung, providing additional evidence for the

[1]Sensory Systems Laboratory, Institute of Zoology, University of Cologne, 50923 Cologne, Germany. [2]Institute of Molecular and Cellular Biosciences, University of Tokyo, Yayoi, Bunkyo-ku, 113-0032 Tokyo, Japan. [3]School of Life Sciences, Tokyo University of Pharmacy and Life Sciences, 1432-1, Horinouchi, Hachioji, 192-0392 Tokyo, Japan. [4]Johann-Friedrich-Blumenbach-Institute, University of Göttingen, 37073 Göttingen, Germany. [5]Theodor-Boveri-Institute, Department of Genetics and Neurobiology, Julius-Maximilians-University of Würzburg, Am Hubland, 97074 Würzburg, Germany. †Present address: Division of Biology 216-76, California Institute of Technology, Pasadena, California 91125, USA.
*These authors contributed equally to this work.

©2009 Macmillan Publishers Limited. All rights reserved

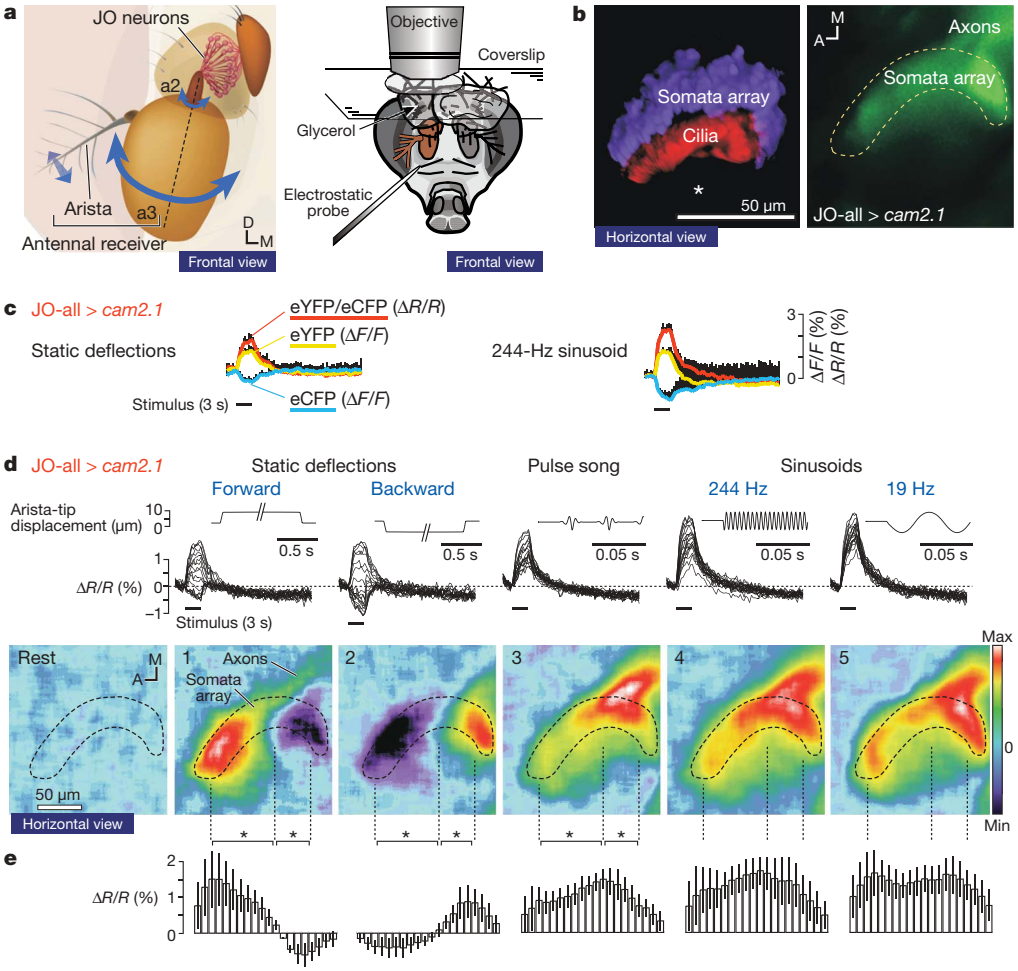


Figure 1 | Mechanically evoked calcium signals in JO neurons. **a**, Antennal anatomy and experimental setup. Left: in response to external forces, the third antennal segment (a3) and the arista twist back and forth (arrows) as a rigid body (antennal receiver), thereby activating JO neurons in the second antennal segment (a2). Right: stimulus forces were imposed on the arista by means of an electrostatic probe, eliciting calcium signals in JO neurons that were monitored with a fluorescence microscope. D, dorsal; M, medial. **b**, Horizontal views of JO. Left: three-dimensional confocal projection. Nuclei and cilia of JO neurons are labelled with anti-Elav antibody (blue) and phalloidin (red). Asterisk: attachment site between JO neurons and a3. Right: cam2.1 fluorescence in a JO-all > *cam2.1* fly as seen through the cuticle of a2. A, anterior. **c**, Time traces of JO calcium signals. Mechanical stimuli evoked reciprocal fluorescent changes ($\Delta F/F$) between eCFP (blue line) and eYFP (yellow line) by FRET. $\Delta R/R$ (%) is the change in eYFP/eCFP fluorescence ratio, where R is the average eYFP/eCFP ratio before stimulus onset and ΔR is the deviation from R (mean and s.d.; $n = 5$ repetitions). Black horizontal bars: stimulus (duration 3 s). **d**, Top: superimposed time traces of responses of JO neurons across the somata array. Insets: arista-tip displacement. Bottom: pseudocoloured ratio changes. $^*P < 0.05$. **e**, Amplitude distribution of ratio changes across the JO somata array (mean ± s.d.; $n = 5$ animals).

functional significance of the measured calcium signals. A small response to static deflection was observed in *nan* mutants (Supplementary Fig. 1b), consistent with the role of Nan in electrical signal propagation rather than transduction suggested in a previous report[23].

Stimulus-specific neural activities in JO

Because the fly's antennal receiver is suspended by a hinge between the second and third segments, it vibrates back and forth in response to acoustic stimuli[12,14] and will deflect backwards and forwards if the fly walks up or down (see Supplementary Information footnotes 1 and 2). By measuring calcium signals in various areas of the JO neuron somata array, we found that deflecting and vibrating the antennal receiver evokes different neural activity patterns in JO (Fig. 1d, e and Supplementary Video 1). When the receiver was deflected statically with a constant force stimulus, opposing calcium signals were seen in the anterior and posterior regions: deflecting the receiver forwards evoked positive signals in the anterior region and negative signals in the posterior one; backward deflection evoked signals of inversed sign (Fig. 1d, e, panels 1 and 2). Broadly distributed signals that peaked in or near the centre region of the somata array, in contrast, were evoked by receiver vibrations induced by recorded courtship songs (pulse song, interpulse interval of ~35 ms or 29 Hz, dominant pulse frequency of ~200 Hz) or sinusoids at high (244 Hz) or low (19 Hz) frequencies (Fig. 1d, e, panels 3–5).

The opposing calcium responses against static deflections are likely to reflect the opposing arrangement of the JO neurons: the fly's JO neurons connect perpendicularly to the anterior and posterior sides of the antennal receiver[12,13,19]. As judged from the anatomy of this connection, deflecting the receiver forwards will stretch JO neurons in the anterior region and compress JO neurons in the posterior. Thus, JO neurons are activated (that is, depolarized) by stretch and deactivated (that is, hyperpolarized) by compression (see Supplementary Information footnote 3 for further discussion).

Vibration- and deflection-sensitive JO neurons

Anatomically, the fly's JO neurons can be subdivided into five subgroups that target distinct zones of the antennal mechanosensory and motor centre (AMMC) in the brain[13] (Fig. 2b, Supplementary Fig. 2 and Supplementary Videos 2 and 3). Each JO neuron typically innervates only one zone of the AMMC, and neurons targeting the same zone

©2009 Macmillan Publishers Limited. All rights reserved

cluster together in JO[13]. To test whether these neural subgroups differ in function, we selectively expressed *cam2.1* using subgroup-specific *GAL4* drivers: JO-B strain for driving expression in JO neuron subgroup B (~100–150 neurons[13]), JO-AB[24] for subgroups A (~50–100 neurons[13]) and B, and JO-CE for subgroups C and E (together ~200 neurons[13]). (Subgroup D, with <30 neurons[13], was not investigated owing to the lack of specific driver lines.)

By using these lines, we found that JO neuron subgroups A and B (AB) and C and E (CE) respond preferentially to different stimulus types: whereas the former were activated maximally by receiver vibrations, the latter responded maximally to static receiver deflections (Fig. 2a, c). The deflection-evoked responses of subgroups CE persisted as long as the deflection was maintained, documenting tonic response characteristics of these neurons (Fig. 2d). The vibration-evoked responses of subgroups AB, in turn, were found to be frequency-dependent (Fig. 2e): when measured in combination, subgroup A and B neurons responded to receiver vibrations at broad frequency ranges between 19 Hz and 952 Hz. When measured alone, however, subgroup B displayed a clear preference for low-frequency vibrations, indicating that subgroup A mainly contributes to the high-frequency responses displayed by the combination of subgroups AB.

JO neurons for gravity sensing

Functional imaging showed that JO neurons of subgroups CE respond preferentially to receiver deflections imposed by static stimuli such as gravitational force (Fig. 2 and Supplementary Fig. 3). To test whether these neurons are required for gravity sensing, we monitored the fly's negative gravitaxis behaviour in a countercurrent apparatus[25]. In this assay, flies are partitioned up into six tubes by giving them the choice five times to stay or to climb up the side of the tube (Fig. 3a and Supplementary Video 4). The partition coefficient Cf describing the final distribution ($0 < Cf < 1$) is large if the flies tend to climb up, and

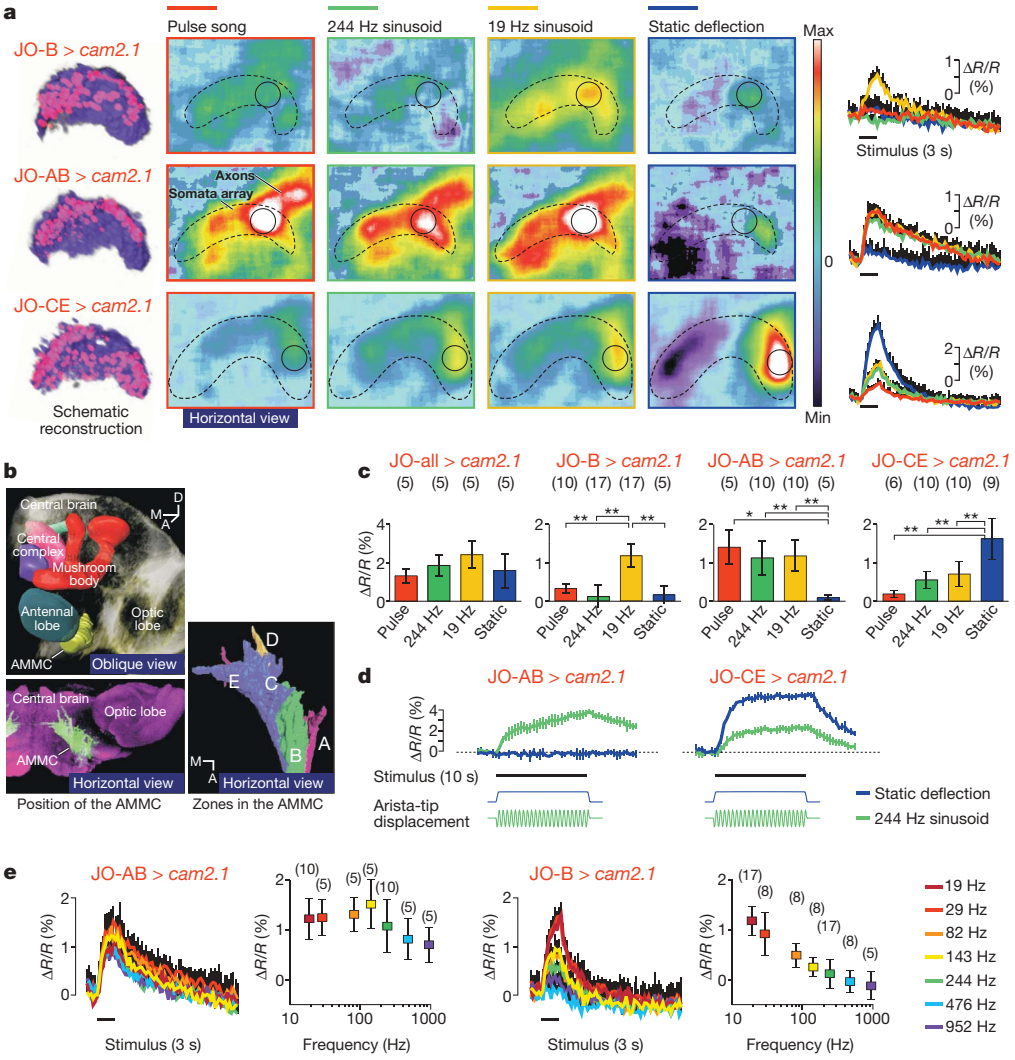


Figure 2 | Responses of JO neuron subgroups. a, Left: schematic horizontal view of the labelled neurons (magenta) and all somata (blue) of JO. Middle: representative pseudocolour images of ratio changes in JO neuron subgroups B, AB and CE evoked by four mechanical stimuli. Right: time traces obtained from the regions of the somata array of labelled neurons that, for given stimuli, showed the largest response (encircled by dashed lines in images, mean and s.d.; $n = 5$ repetitions). Note that absolute signal amplitudes may differ between fly strains owing to differences in labelled neuron numbers and expression levels. **b**, Architecture of the AMMC. Top left: location of the AMMC in the fly brain (schematic three-dimensional reconstruction of confocal serial sections). Bottom left: horizontal view of the brain at the level of the AMMC. Right: target zones of JO neuron subgroups in the AMMC. **c**, Average ratio change in the somata region that showed the largest response (mean ± s.d.; numbers in parentheses represent the number of animals). Subgroups AB respond preferentially to vibrations, and subgroups CE to static deflections. $^{*}P < 0.05$; $^{**}P < 0.01$. **d**, Superimposed time traces of the ratio changes to long stimuli (10 s, black horizontal bar) in subgroups AB (left) and CE (right) (mean ± s.d.; $n = 5$ repetitions). **e**, Time traces of ratio changes (mean ± s.d.; $n = 5$ repetitions, black horizontal bars indicate stimulus duration) and averaged ratio changes (mean ± s.d.; numbers in parentheses represent number of animals) in JO neuron subgroups AB (left two panels) and B (right two panels) measured at 19–952 Hz.

©2009 Macmillan Publishers Limited. All rights reserved

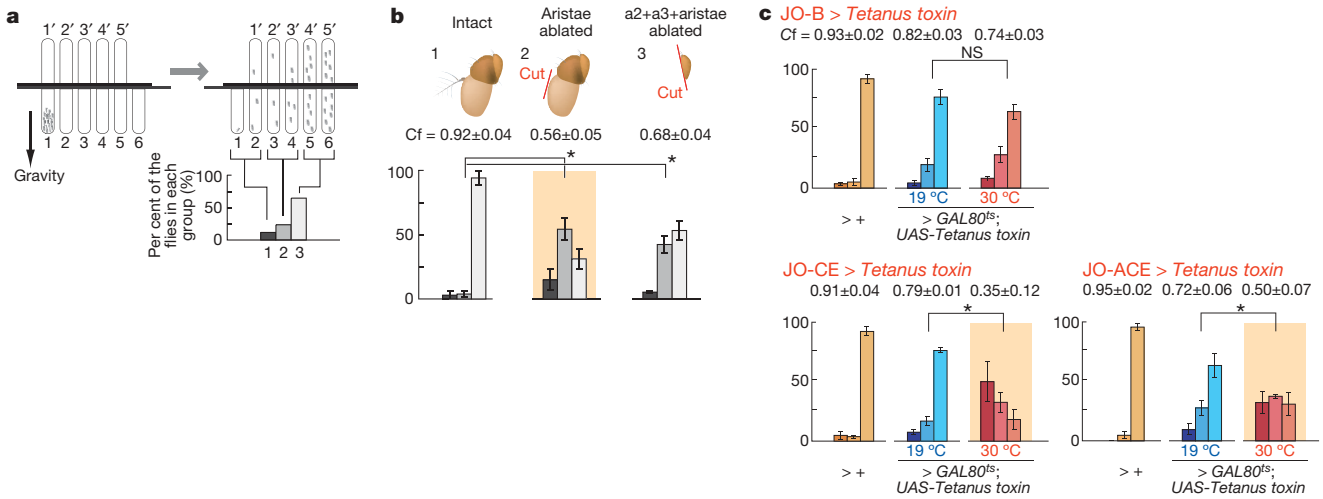


Figure 3 | Requirement of JO neuron subgroups for gravity detection. **a**, Negative gravitaxis assay. Numbers 1 to 6 and 1′ to 5′ represent the lower and upper tubes, respectively, from left to right. **b**, Wild-type flies with intact and ablated antennae. Cf, partition coefficient of the final distribution (mean ± s.e.m., >5 trials for each experimental group, see Supplementary Information footnotes 4 and 5); *$P < 0.05$, Student's *t*-test. **c**, Flies with genetically silenced JO neurons (mean ± s.e.m., >5 trials for each experimental group). *$P < 0.05$, Student's *t*-test. Cases with aberrant behaviour are highlighted (Cf < 2/3; see Supplementary Information footnote 5). Negative gravitaxis is eliminated by silencing subgroups CE and ACE, but not B. NS, not significant. The *x* and *y* axes for **b** and **c** are the same as in **a**.

small if they tend to stay (see Supplementary Information footnotes 4 and 5). As expected, wild-type flies displayed negative gravitaxis behaviour (Fig. 3b). This behaviour, but not phototaxis (Supplementary Fig. 4a), was abolished when the antennal aristae were ablated (Fig. 3b, panel 2). Removing also the third and second antennal segments, the latter of which houses JO, yielded slightly higher Cf values (Fig. 3b, panel 3, $P < 0.1$ between panels 2 and 3). Apparently, when JO is lost, other sense organs may partially take over gravity sensing, for example, receptors on the neck and legs that have been implicated in gravity sensing in other insect species[2,26].

To silence selectively subgroups of JO neurons, we conditionally expressed tetanus toxin[27] using subgroup-specific *GAL4* drivers and *tubulin-GAL80^{ts}*, a temperature-sensitive blocker of Gal4 expressed ubiquitously by the tubulin promoter[28,29]. Tetanus toxin expression was activated shortly before behavioural experiments by raising the rearing temperature from 19 °C to 30 °C. Expressing tetanus toxin by means of JO-all and JO-AB *GAL4* drivers caused general locomotion defects as indicated by aberrant phototaxis, probably due to Gal4 expression elsewhere in the body (Supplementary Fig. 4b). When tetanus toxin was expressed by means of the drivers JO-B, JO-CE and JO-ACE, however, phototaxis was normal (Supplementary Fig. 4c). Using these lines, we found that silencing subgroups CE and ACE, but not subgroup B, abolishes gravitaxis (Fig. 3c). Hence, consistent with the physiological data, the fly's gravitaxis behaviour requires the deflection-sensitive JO neurons of subgroups CE.

Vibration-responsive neurons are required for hearing

To determine which JO neurons are required for hearing, we next exposed groups of males to synthesized pulse-song of increasing intensity. This made wild-type males chase other males to form courtship chains[4] (Fig. 4a and Supplementary Video 5). Consistent with earlier reports[4], we found that ablating the distal antennal segments abolishes this sound-evoked behaviour (Fig. 4b, c). We further found that this behaviour specifically requires JO neurons of subgroup B: whereas expressing tetanus toxin in subgroup B impaired the male's chaining behaviour, the behaviour remained unaffected when tetanus toxin was targeted to subgroups CE or ACE (Fig. 4d and Supplementary Fig. 4d).

Although physiological data indicate a role of subgroup-A JO neurons in sound detection (Fig. 2), silencing these neurons did not affect responses to courtship song (Fig. 4d). One possible explanation is that the JO-ACE driver used in the behavioural experiments labels a fraction of subgroup-A neurons[13]; not all subgroup-A neurons were therefore silenced by tetanus toxin. Additional hints on solving the apparent conundrum were obtained when we investigated how ablating specific subgroups affects sound-evoked compound action potentials (CAPs; the sum of action potentials recorded extracellularly) in the antennal nerve[18,19]. We induced selective apoptosis by expressing ricin toxin A[30] under Gal4 control using the *eyFLP/FRT* system[31], which drives expression of flippase (FLP) enzyme by the enhancer fragment of *eyeless* (*ey*) gene. FLP induces recombination, which leads to the removal of a stop between two *FRT* sites to restrict *ricin toxin* expression to *GAL4*-expressing cells in the eye and antenna (Supplementary Fig. 5a–d). We then sinusoidally vibrated the antennal receiver while simultaneously monitoring the arista's displacement and the CAPs in the nerve. The amplitude of the CAP increased sigmoidally for the antennal displacement range of ~25 nm–1 μm in wild-type flies as well as in the flies in which JO neuron subgroups B or BCE were ablated (Fig. 4e), independent of the frequency of stimulation (Supplementary Fig. 5e), but the range shifted up to ~100 nm–4 μm when also subgroup A was ablated (Fig. 4e and Supplementary Fig. 5f). Hence, subgroup A is probably required for the detection of nanometre-range receiver vibrations as imposed by attenuated pulse-songs and/or the faint sine-songs of courting males[5].

NompC is expressed in sound-sensitive neurons

To gain first insights into the molecular mechanisms that account for the functional differences between deflection- and vibration-sensitive JO neurons, we analysed which JO neurons express the candidate mechanotransducer channel NompC (no mechanoreceptor potential C, also known as TRPN1)[23,32]. To identify *nompC*-expressing neurons, we expressed *GAL4* under the control of the *nompC* promoter (*nompC-GAL4*)[33]. In contrast to *F-GAL4*, which expresses Gal4 under the control of the *nanchung* promoter and labels almost all JO neurons, only some JO neurons were labelled by *nompC-GAL4* (Fig. 5a). Projection analysis revealed that *nompC-GAL4* labels JO neurons of subgroups AB but not CE (Fig. 5b and Supplementary Fig. 6b). Hence, whereas the TRPV channel Nanchung is expressed by almost all JO neurons, the TRPN channel NompC seems specific for sound-sensitive JO neurons. This differential expression presumably explains why disrupting NompC reduces, but does not abolish, mechanically evoked responses in the fly's antennal nerve[34], supporting NompC as a candidate mechanotransducer for hearing and indicating that gravity transduction is independent of NompC.

©2009 Macmillan Publishers Limited. All rights reserved

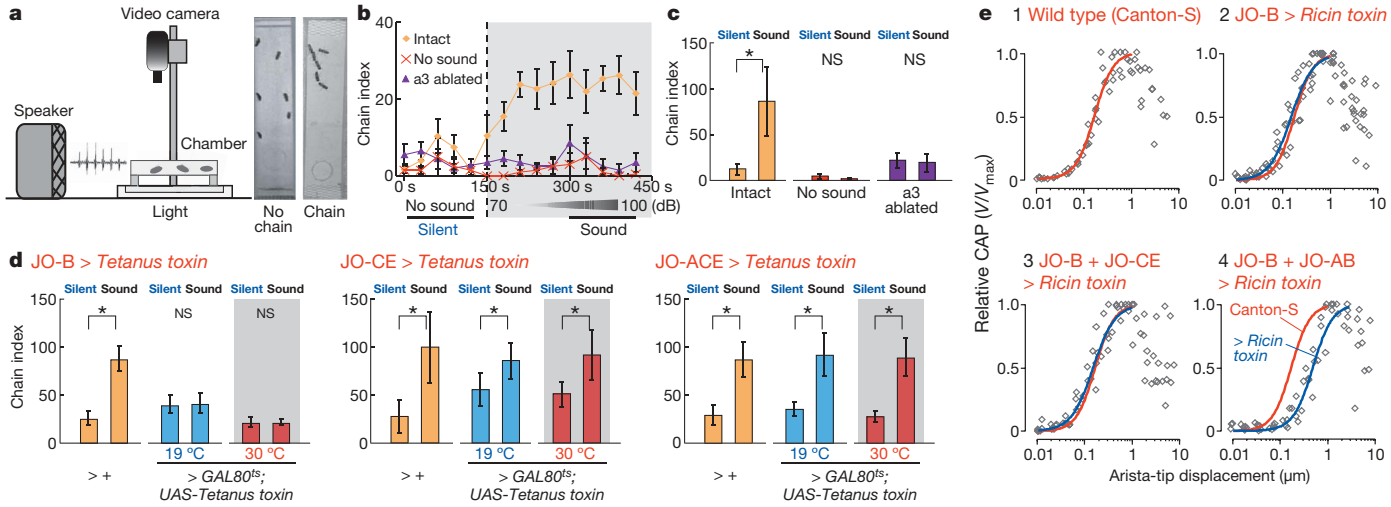


Figure 4 | Requirement of JO neuron subgroups for hearing. **a**, Courtship-song-detection assay. Representative images taken from videos of flies forming a courtship chain are shown. **b**, Time course of chain formation. Wild-type flies form chains only if the antenna is intact and in the presence of sound. **c**, Chain indices for flies in **b**. **d**, Chain indices for flies with genetically silenced JO neurons (mean ± s.e.m., >5 trials for each experimental group). Grey boxes: neurons silenced. Silencing subgroup B, but not subgroups CE and ACE, eliminates the formation of courtship chains. JO-B > *Tetanus toxin* flies fail to form chains even at 19 °C, probably due to leaky GAL80^{ts} suppression. *$P < 0.05$; NS: $P > 0.05$, Mann–Whitney U-tests. **e**, Amplitude of sound-evoked CAPs in the antennal nerve as function of arista-tip displacement in ricin-toxin-expressing flies and controls. In panels 3 and 4, two *GAL4* driver lines were crossed to ablate larger cell populations. Blue, average fits for each genotypes; red, repeated in each panel, average fit for wild-type controls.

Central circuits for gravity and sound

As judged from their central projections, gravity- and sound-sensitive JO neurons target distinct primary centres in the AMMC and feed into distinct brain circuits. To trace these circuits, we screened 3,939 *GAL4* enhancer trap lines[35] for higher-order neurons in the *Drosophila* brain that arborize in the AMMC. The target zones of subgroups A and B in the AMMC, which form the primary auditory centres, are both characterized by a close association with the inferior part of the ventrolateral protocerebrum (VLP), which is also directly supplied by a subset of subgroup-A neurons[13] and can be regarded as the secondary auditory centre: various interneurons were identified that arborize in both the VLP and the target zones of subgroups AB in the AMMC (Fig. 6a and Supplementary Fig. 6a, see also Supplementary Information footnote 6). These zones are also characterized by extensive commissural connections, with interneurons connecting the contralateral zones by means of commissures above and below the oesophagus (Fig. 6a and Supplementary Fig. 7). Also the giant fibre neuron (GFN), a large descending neuron that controls jump escape behaviour[36,37], arborizes in zone A and in the inferior VLP (Fig. 6a, see also Supplementary Information footnote 7). The GFNs of both sides are connected by means of the giant commissural interneurons[37], a feature not observed in the other descending neurons described below. All higher-order neurons we identified arborized only in the target zone of either subgroup A or B, pointing to a parallel organization of the auditory pathway that might explain why silencing only one subgroup of vibration-sensitive neurons suffices to abolish the flies' sound-evoked behaviour.

Aside from a few JO neurons of subgroups CE that directly cross the midline[13], we did not find commissural connections between the target zones of subgroups CE (Supplementary Fig. 7). No connections

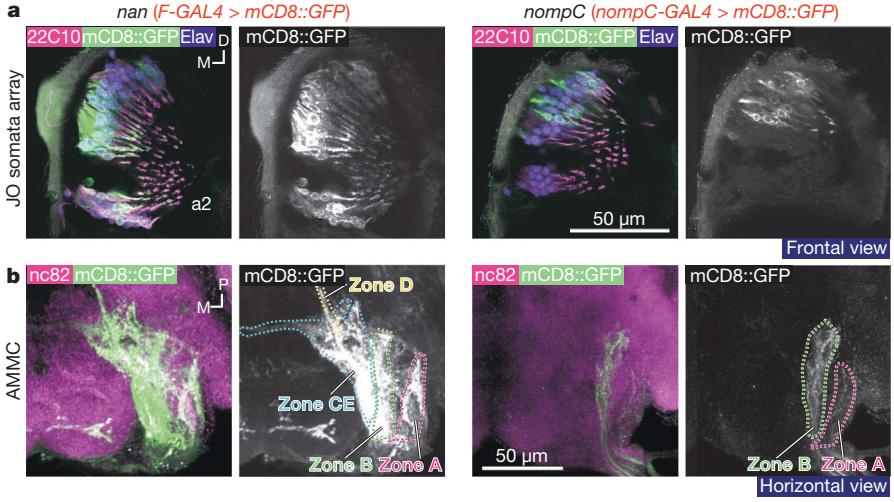


Figure 5 | Expression of *nan* and *nompC*. **a**, Distribution of *nan*- and *nompC*-expressing neurons in JO (confocal projections). JO neurons are visualized by 22C10 antibody (magenta) and a pan-neural marker, anti-Elav antibody (blue). Left: *F-GAL4*, in which *GAL4* is fused to the *nan* promoter, drives the expression of mCD8::GFP reporter proteins (green) in virtually all JO neurons. Right: only a subset of JO neurons is labelled if *GAL4* is fused to the promoter of *nompC*. **b**, Confocal projection images of the brain counter-labelled with presynaptic antibody nc82 (magenta). *eyFLP* was used to restrict Gal4-mediated *GFP* expression to the eye and antenna. *F-GAL4* labels JO neurons innervating all zones of the AMMC (left), whereas *nompC-GAL4* labels subgroup-B neurons and a subset of subgroup A, but not subgroups CE (right).

©2009 Macmillan Publishers Limited. All rights reserved

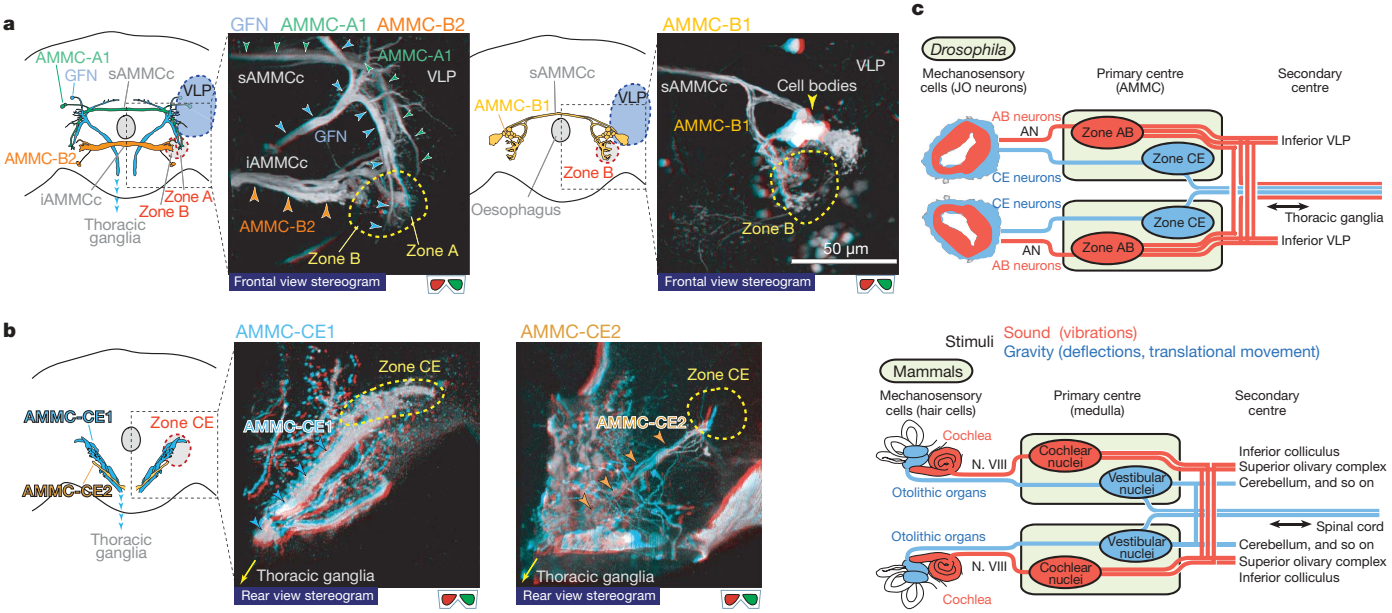


Figure 6 | Higher-order neurons in the AMMC. a, Diagrams (left) and three-dimensional confocal projection stereograms (right, for red/green glasses) of higher-order neurons arborizing in the target zones of subgroups AB. AMMC zones and the VLP are highlighted. Arrowheads point to AMMC-A1, -B1 and -B2 neurons and the GFN. The neural pathway downstream of subgroups-AB neurons displays a secondary centre in the inferior VLP and commissural connections between hemispheres. sAMMCc/iAMMCc, superior/inferior AMMC commissures (commissures above and below the oesophagus connecting AMMCs). **b**, Diagram and three-dimensional confocal projection images of higher order neurons arborizing in the target zones of subgroups CE. Arrowheads indicate the structure of AMMC-CE1 and -CE2 neurons, respectively. Subgroups CE neurons have direct connections with the descending tracts to the thoracic ganglia. **c**, Schematic comparison of mechanosensory pathways in flies and mammals. AN, antennal nerve; N. VIII, eighth cranial nerve. For details, see Supplementary Information footnote 8.

between these zones and the VLP were identified either. These zones, however, were abundantly contributed to by descending and ascending neurons to and from the thoracic ganglia (Fig. 6b and Supplementary Fig. 6a). Together, the tight commissural connection in the pathways downstream of sound-sensitive JO neurons and abundant descending tracts downstream of gravity-sensitive JO neurons are reminiscent of the connectivities of mammalian auditory and vestibular pathways (Fig. 6c), the former of which has extensive binaural interactions between the secondary centres of both hemispheres[6,38] whereas the latter has direct descending pathways from the primary centre to the spinal cord[7,39,40] (for more detail, see Supplementary Information footnote 8).

Discussion

Housing almost 480 primary mechanosensory neurons[13], JO is the largest mechanosensory organ of the fruitfly. We have shown that this organ serves at least two mechanosensory submodalities that are segregated at the level of the primary neurons. JO neurons of subgroups AB respond preferentially to antennal vibrations; they differ in their frequency characteristics, express the NompC channel, and have a role in sound detection. JO neurons of subgroups CE respond preferentially to static deflections, provide information about the forcing direction, do not express the NompC channel, and are required for gravity sensing. As judged from our imaging data and antennal nerve recordings, JO neurons of subgroups CE respond to tiny displacements imposed by the Earth's gravitational field (see Supplementary Information footnote 1 and Supplementary Fig. 3a). Subgroups-CE neurons also respond to large antennal displacements as may be imposed by air jets or wind (see accompanying manuscript[41], Supplementary Information footnote 9 and Supplementary Fig. 3c), indicating either that the same subgroups-CE neurons mediate gravity and wind detection or, alternatively, that sensitive, gravity-responsive CE neurons and less-sensitive, wind-responsive CE neurons may coexist.

As all JO neurons attach to the same antennal receiver, how do their distinct response characteristics come about? The opposing calcium signals evoked by receiver deflections are likely to reflect the opposing connections of JO neurons with the antennal receiver[12,13,19], indicating that these neurons are hyperpolarized by compression and depolarized by stretch (see Supplementary Information footnote 3). The vibration- and deflection-sensitivities of distinct JO neuron subgroups may reflect differences in the molecular machineries for transduction; JO neurons reportedly harbour adapting channels that transduce dynamic receiver vibrations but fully adapt within milliseconds during static receiver deflection[17,19]. Because deflecting the receiver statically for several seconds evokes sustained large-amplitude calcium signals in subgroups CE (Fig. 2a, d), however, also less- or non-adapting channels seem to exist. Transduction channels with different adaptation characteristics seem to occur in many mechanosensory systems, including the mammalian cochlea[42] and also *Drosophila* bristle neurons, which reportedly display mechanically evoked adapting, NompC-dependent and also non-adapting, NompC-independent currents[32]. In the fly's JO, such functional and molecular specializations of the transduction machineries could explain why some neurons preferentially respond to gravity whereas others preferentially respond to sound. The segregation of gravitational and auditory stimuli in the *Drosophila* JO may thus take place at the very first stage of neuronal signal processing.

METHODS SUMMARY

See Supplementary Information footnote 10 for fly genotypes.

Stimulation. The antennal receiver was actuated by feeding voltage commands to an external electrode that served as an electrostatic probe[17]. To allow for attractive and repulsive forcing, the potential of the fly's body was lowered to −15 V against ground[17]. Voltage-force characteristics were flat for frequencies <5 kHz. Acoustic stimuli were used for behavioural and CAP assays. For the equivalence of acoustically and electrostatically induced receiver movements, see ref. 17.

Calcium imaging. Fluorescence signals were monitored using a CCD camera (CoolSnap HQ, Roper Scientific) mounted on a microscope (Axioscop2, Carl Zeiss)[22] (also A.K., T.E., M.C.G. and A.F., manuscript in preparation). Each experiment was performed in ≥5 flies. Responses to five repetitive stimuli were averaged. Data acquisition and evaluation were performed as described[22].

Receiver displacements. Displacements were measured at the tip of the arista using a Polytec PSV-400 laser Doppler vibrometer[17,18]. In fly strains used for

©2009 Macmillan Publishers Limited. All rights reserved

imaging, receiver fluctuations support the integrity of the antenna and JO neurons[18] (Supplementary Table 1).

Behavioural assays. Sound and gravity responses were assayed as described[3,4] (also H.K.I., A.K. and K.I., manuscript in preparation). Between 30 and 50 flies were used for each experiment. Sound detection was examined in six males at a time. To produce intensity profiles (Fig. 4b), males forming courtship chains were scored each 3 s and summed up for 30 s (maximum chain index of 60). For comparisons between flies under silent and sound-stimulated conditions (Fig. 4c, d), scores were summed for 150 s (maximum chain index of 300). For statistical analyses, see Supplementary Information footnotes 4 and 5.

Nerve recordings. CAP responses were recorded by means of a tungsten electrode inserted between the antenna and the head. The indifferent electrode was inserted into the thorax. For each genotype, $\geq$7 flies were examined.

Neuroanatomy. Serial optical sections of adult fly brains and antennae were captured using confocal microscopes and three-dimensionally reconstructed as described[13]. See Supplementary Methods for detailed equipments.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 27 June 2008; accepted 20 January 2009.

1. Toma, D. P., White, K. P., Hirsch, J. & Greenspan, R. J. Identification of genes involved in *Drosophila melanogaster* geotaxis, a complex behavioral trait. *Nature Genet.* **31,** 349–353 (2002).
2. Beckingham, K. M., Texada, M. J., Baker, D. A., Munjaal, R. & Armstrong, J. D. Genetics of graviperception in animals. *Adv. Genet.* **55,** 105–145 (2005).
3. Tempel, B. L., Livingstone, M. S. & Quinn, W. G. Mutations in the dopa decarboxylase gene affect learning in *Drosophila. Proc. Natl Acad. Sci. USA* **81,** 3577–3581 (1984).
4. Eberl, D. F., Duyk, G. M. & Perrimon, N. A genetic screen for mutations that disrupt an auditory response in *Drosophila melanogaster. Proc. Natl Acad. Sci. USA* **94,** 14837–14842 (1997).
5. Tauber, E. & Eberl, D. F. Acoustic communication in *Drosophila. Behav. Processes* **64,** 197–210 (2003).
6. Hudspeth, A. J. in *Principles of Neural Science* (eds Kandel, E. R., Schwartz, J. H. & Thomas, M. J.) 590–613 (McGraw-Hill, 2000).
7. Goldberg, M. E. & Hudspeth, A. J. in *Principles of Neural Science* (eds Kandel, E. R., Schwartz, J. H. & Thomas, M. J.) 801–815 (McGraw-Hill, 2000).
8. Todi, S. V., Sharma, Y. & Eberl, D. F. Anatomical and molecular design of the *Drosophila* antenna as a flagellar auditory organ. *Microsc. Res. Tech.* **63,** 388–389 (2004).
9. Kim, J. *et al.* A TRPV family ion channel required for hearing in *Drosophila. Nature* **424,** 81–84 (2003).
10. Caldwell, J. C. & Eberl, D. F. Towards a molecular understanding of *Drosophila* hearing. *J. Neurobiol.* **53,** 172–189 (2002).
11. Kernan, M. J. Mechanotransduction and auditory transduction in *Drosophila. Pflugers Arch.* **454,** 703–720 (2007).
12. Göpfert, M. C. & Robert, D. The mechanical basis of *Drosophila* audition. *J. Exp. Biol.* **205,** 1199–1208 (2002).
13. Kamikouchi, A., Shimada, T. & Ito, K. Comprehensive classification of the auditory sensory projections in the brain of the fruit fly *Drosophila melanogaster. J. Comp. Neurol.* **499,** 317–356 (2006).
14. Göpfert, M. C. & Robert, D. Biomechanics. Turning the key on *Drosophila* audition. *Nature* **411,** 908 (2001).
15. Baker, D. A., Beckingham, K. M. & Armstrong, J. D. Functional dissection of the neural substrates for gravitaxic maze behavior in *Drosophila melanogaster. J. Comp. Neurol.* **501,** 756–764 (2007).
16. Dickson, B. J. Wired for sex: the neurobiology of *Drosophila* mating decisions. *Science* **322,** 904–909 (2008).
17. Albert, J. T., Nadrowski, B. & Göpfert, M. C. Mechanical signatures of transducer gating in the *Drosophila* ear. *Curr. Biol.* **17,** 1000–1006 (2007).
18. Albert, J. T., Nadrowski, B., Kamikouchi, A. & Göpfert, M. C. Mechanical tracing of protein function in the *Drosophila* ear. *Nature Protocols*. doi:10.1038/nprot.2006.364 (2006).
19. Nadrowski, B., Albert, J. T. & Gopfert, M. C. Transducer-based force generation explains active process in *Drosophila* hearing. *Curr. Biol.* **18,** 1365–1372 (2008).
20. Brand, A. H. & Perrimon, N. Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. *Development* **118,** 401–415 (1993).
21. Miyawaki, A., Griesbeck, O., Heim, R. & Tsien, R. Y. Dynamic and quantitative Ca^{2+} measurements using improved cameleons. *Proc. Natl Acad. Sci. USA* **96,** 2135–2140 (1999).
22. Fiala, A. & Spall, T. *In vivo* calcium imaging of brain activity in *Drosophila* by transgenic cameleon expression. *Sci. STKE* **2003,** pl6 (2003).
23. Göpfert, M. C., Albert, J. T., Nadrowski, B. & Kamikouchi, A. Specification of auditory sensitivity by *Drosophila* TRP channels. *Nature Neurosci.* **9,** 999–1000 (2006).
24. Sharma, Y., Cheung, U., Larsen, E. W. & Eberl, D. F. PPTGAL, a convenient Gal4 P-element vector for testing expression of enhancer fragments in *Drosophila. Genesis* **34,** 115–118 (2002).
25. Benzer, S. Behavioral mutants of *Drosophila* isolated by countercurrent distribution. *Proc. Natl Acad. Sci. USA* **58,** 1112–1119 (1967).
26. Horn, E. & Lang, H.-G. Positional head reflexes and the role of the prosternal organ in the walking fly, *Calliphora erythrocephala. J. Comp. Physiol. [A]* **126,** 137–146 (1978).
27. Sweeney, S. T., Broadie, K., Keane, J., Niemann, H. & O'Kane, C. J. Targeted expression of tetanus toxin light chain in *Drosophila* specifically eliminates synaptic transmission and causes behavioral defects. *Neuron* **14,** 341–351 (1995).
28. McGuire, S. E., Le, P. T., Osborn, A. J., Matsumoto, K. & Davis, R. L. Spatiotemporal rescue of memory dysfunction in *Drosophila. Science* **302,** 1765–1768 (2003).
29. Thum, A. S. *et al.* Differential potencies of effector genes in adult *Drosophila. J. Comp. Neurol.* **498,** 194–203 (2006).
30. Smith, H. K. *et al.* Inducible ternary control of transgene expression and cell ablation in *Drosophila. Dev. Genes Evol.* **206,** 14–24 (1996).
31. Newsome, T. P., Asling, B. & Dickson, B. J. Analysis of *Drosophila* photoreceptor axon guidance in eye-specific mosaics. *Development* **127,** 851–860 (2000).
32. Walker, R. G., Willingham, A. T. & Zuker, C. S. A *Drosophila* mechanosensory transduction channel. *Science* **287,** 2229–2234 (2000).
33. Liu, L. *et al. Drosophila* hygrosensation requires the TRP channels *water witch* and *nanchung. Nature* **450,** 294–298 (2007).
34. Eberl, D. F., Hardy, R. W. & Kernan, M. J. Genetically similar transduction mechanisms for touch and hearing in *Drosophila. J. Neurosci.* **20,** 5981–5988 (2000).
35. Otsuna, H. & Ito, K. Systematic analysis of the visual projection neurons of *Drosophila melanogaster*. I. Lobula-specific pathways. *J. Comp. Neurol.* **497,** 928–958 (2006).
36. Bacon, J. P. & Strausfeld, N. J. The dipteran 'Giant fibre' pathway: neurons and signals. *J. Comp. Physiol. [A]* **158,** 529–548 (1986).
37. Phelan, P. *et al.* Mutations in *shaking-B* prevent electrical synapse formation in the *Drosophila* giant fiber system. *J. Neurosci.* **16,** 1101–1113 (1996).
38. Cant, N. B. & Benson, C. G. Parallel auditory pathways: projection patterns of the different neuronal populations in the dorsal and ventral cochlear nuclei. *Brain Res. Bull.* **60,** 457–474 (2003).
39. Barmack, N. H. Central vestibular system: vestibular nuclei and posterior cerebellum. *Brain Res. Bull.* **60,** 511–541 (2003).
40. Büttner-Ennever, J. A. A review of otolith pathways to brainstem and cerebellum. *Ann. NY Acad. Sci.* **871,** 51–64 (1999).
41. Yorozu, S. *et al.* Distinct sensory representations of wind and near-field sound in the *Drosophila* brain. *Nature* doi:10.1038/nature07843 (this issue).
42. Ricci, A. J., Kennedy, H. J., Crawford, A. C. & Fettiplace, R. The transduction channel filter in auditory hair cells. *J. Neurosci.* **25,** 7831–7839 (2005).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank D. F. Eberl for JO15, C. J. O'Kane for UFWTRA19, B. J. Dickson for *UAS-GFP S65T* and *eyFLP* fly strains, H. Tanimoto for flies carrying *tubulin-GAL80^{ts}* and *UAS-tetanus toxin*, C. Kim for *nan^{dy5}*, M. J. Kernan for *nan^{36a}*, L. Liu for *nompC-GAL4.25*, A. Wong and G. Struhl for *UAS > CD2, y > CD8::GFP*, J. Urban and G. Technau for MZ-series enhancer trap strains, the members of the NP consortium (a group of eight laboratories in Japan that together produced a large collection of *GAL4* lines) and D. Yamamoto for the NP-series strains, Bloomington Stock Centre for *elav^{c155}-GAL4*, D. F. Eberl and C. P. Kyriacou for courtship sound data, S. Fujita for 22C10 antibody, the Developmental Studies Hybridoma Bank for antibodies anti-Elav and nc82, T. Völler for help with calcium imaging, H. Otsuna and K. Shinomiya for preparing some figures, M. Dübbert, K. Öchsner, M. Matsukuma, S. Shuto and K. Yamashita for technical assistance, J. T. Albert, E. D. Hoopfer, B. Nadrowski, K. Endo, H. Otsuna, Y. Hiromi, E. Buchner and N. J. Strausfeld for discussion, and D. J. Anderson and S. Yorozu for sharing unpublished data. This work was supported by the Japanese Cell Science Research Foundation, the Alexander von Humboldt Foundation, and the Japan Society for the Promotion of Science (to A.K.), the DFG Collaborative Research Centre 554 (to A.F.), the Volkswagen Foundation, the BMBF Bernstein Network for Computational Neuroscience, and the DFG Research Centre Molecular Physiology of the Brain (to M.C.G.), and the Human Frontier Science Program Organisation, BIRD/Japan Science and Technology Agency, and the Japan Society for the Promotion of Science (to K.I.).

Author Contributions A.K., M.C.G. and K.I. designed research; A.K. and A.F. performed calcium imaging. A.K. and H.K.I. performed fly genetics; H.K.I. performed behavioural and anatomical experiments; T.E. performed nerve recordings; A.K., H.K.I. and O.H. performed histology; A.K., H.K.I., M.C.G. and K.I. wrote the paper; and M.C.G. and K.I. supervised the work. All authors discussed the concepts and results, and commented on the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to K.I. (itokei@iam.u-tokyo.ac.jp) or M.C.G. (mgoepfe@gwdg.de).

©2009 Macmillan Publishers Limited. All rights reserved

METHODS

Fly stocks. The following *GAL4* strains were used: JO-all (*F-GAL4*; ref. 9), JO-B (JO2, also known as NP1046; ref. 13), JO-AB (JO15; ref. 24), JO-CE (JO31, also known as NP6250; ref. 13), JO-ACE (JO4, also known as NP6303; ref. 13) and *nompC-GAL4.25* (ref. 33); other strains included *UAS-GFP S65T* (T2 strain) for visualization, *UAS-cam2.1* (*UAS-cameleon2.1-82* (ref. 22) and *UAS-cameleon2.1-76* (ref. 22)) for calcium imaging, *UAS-tetanus toxin* (ref. 27) and *tubulin-GAL80^{ts}* (refs 28 and 29) for the selective silencing of neurons, *eyFLP* (ref. 31) and UFWTRA19(ref. 30) for ricin-mediated cell ablation, and *eyFLP* and *UAS > CD2, y > CD8::GFP* (ref. 43) for visualizing neurons from the antenna. To visualize neurons downstream of JO neurons, we screened NP- and MZ-series *GAL4* enhancer trap lines[35]. The Canton-S strain was used as the wild type.

Calcium imaging. To enhance reporter signals, flies were made homozygous for both *GAL4* and *UAS-cam2.1*. Only JO-AB was analysed in the heterozygous condition (JO-AB/TM6B) because the antennal mechanics were significantly altered in homozygous JO-AB flies (Supplementary Table 1). After raising flies at 29 °C for 3–15 days to enhance *cam2.1* expression, flies were anaesthetized on ice and affixed onto a coverslip with beeswax. The dorsal tip of the second antennal segment was attached to the coverslip with dental glue, and the gap between the antennae and the coverslip was filled with glycerol. Binning of the cooled CCD camera (CoolSnap HQ, Roper Scientific) was set to give a resolution of 0.645 μm per pixel. A water-immersion ×40 objective (NA = 0.8) was used for imaging. Individual flies were assayed for up to 30 min, with inter-stimulus intervals of 30–60 s. The fluorescence of eCFP and eYFP were captured simultaneously at a rate of 3 Hz with an exposure time of 200 ms. As judged from the mechanical fluctuations of the antennae, the preparation was stable for about 2 h. Only flies with receiver fluctuations indistinguishable from those of the wild-type flies were used for data collection. Receiver displacements used to evoke activities in JO ranged between <5, 10 and 100 μm. Smaller displacements (~1–5 μm) evoked essentially similar response patterns in JO, although with a lower signal-to-noise ratio. Data were analysed off-line with MetaMorph software (Molecular Devices) as described previously[22]. Twenty regions of interest (20 pixels in diameter) were used for analysis, whereby the intensities of eYFP and eCFP fluorescence were normalized to those preceding the stimulus onset ($t = 0$). To compare changes in the eYFP/eCFP ratio across experiments and animals, the mean ratio change at the end of the stimulus was used. Two-tailed Mann–Whitney U-tests were used for statistical analysis because the ratio changes typically did not display a Gaussian distribution. For multiple comparisons, the Sidak–Bonferroni correction was applied[44]. For computing pseudocolour coded ratio changes, the mean of the ratio during the second preceding stimulus onset was subtracted from that during the second preceding the stimulus end.

Behavioural assay. Flies were raised on standard medium in a 12 h light/dark cycle at 19 °C to prevent leaky inactivation of *GAL80^{ts}* (H.K.I., A.K. and K.I., manuscript in preparation). Flies were collected under ice anaesthesia on the day after eclosion. Wings (song-detection assays) or aristae/antennae (ablation experiments) were removed with fine forceps. The flies were then kept at 19 °C. Negative gravitaxis/phototaxis and sound responses were assayed using 3–7- and 10–14-day-old flies, respectively (when raised at 25 °C, this would correspond to ages of roughly 2–4 and 5–7days). To remove the *GAL80^{ts}*-mediated suppression of effecter gene expression, flies were transferred to 30 °C for 24 h before the experiment, and placed back at 19 °C 1 h before the experiment was performed. All assays were carried out at 23–25 °C and 40–60% humidity. Using a countercurrent apparatus[25], we measured startle-induced negative gravitaxis and phototaxis[3]. Gravitaxis was monitored in pitch darkness. Phototaxis was induced by a 40-W fluorescent lamp positioned 30 cm above the centre of the countercurrent apparatus. In brief, we collected flies at the bottom of the tubes by tapping the countercurrent apparatus on the table, and then kept the apparatus still for 30 s to allow the flies to climb the wall. To test the fly's physical ability for climbing the tube wall, phototaxis assay was also performed in a vertical orientation. Five repetitive procedures distributed the flies into six tubes depending on the partition coefficient Cf (that is, their probability of climbing the tubes at each test), which equals mean ± s.e.m. of the weighted mean of the fly numbers in the tubes (see Supplementary Information footnote 4). The Cf value was evaluated as described in Supplementary Information footnote 5. For the sound-response assay, synthesized courtship song[4] was broadcast via a speaker (25 cm in diameter, TAMON S25 W027), with the cone of the speaker being 10 cm away from the centre of the chamber. Behaviour was monitored with a video camera (US522, Panasonic) mounted above the chamber. Recorded movies were converted to serial frames every three seconds, and the number of flies in chains was counted blindly as described[4].

Evaluation of CAP responses. CAPs are the summed action potentials that can be recorded extracellularly from the antennal nerve. CAP responses and antennal displacement data were subjected to fast Fourier transforms (FFT) with a resolution of 1 Hz. CAP responses were quantified by measuring their FFT amplitudes at twice the stimulus frequency, because previous observations had shown the frequency doubling of CAP responses[17,34]. Data analysis and statistical data evaluation were performed using Spike 2 (Cambridge Electronic Design), Excel 2004 (Microsoft) and Sigma Plot 10 (Systat Software). Fits were run with a Hill equation consisting of four parameters.

Immunolabelling of the fly brains and antennae. Adult brains and antennae were dissected from the progeny of *GAL4* strains and *UAS-GFP S65T* (T2) or *eyFLP*; *UAS > CD2, y > CD8::GFP* crosses and labelled as described previously[13]. The antibodies used were: rabbit anti-GFP polyclonal serum (1:300, Invitrogen) and mouse monoclonal antibodies nc82 (1:20, Developmental Studies Hybridoma Bank), 22C10 (1:50, gift from S. Fujita) and rat polyclonal antibody anti-Elav (1:250, Developmental Studies Hybridoma Bank) for primary antibodies, and Alexa Fluor 488 goat anti-rabbit IgG (1:300, Invitrogen), Alexa Fluor 568 goat anti-mouse IgG (1:300, Invitrogen) and Alexa Flour 633 goat anti-mouse IgG (1:300, Invitrogen) for secondary antibodies. Incubations with primary and secondary antibodies were 72 h and 48 h, respectively.

43. Wong, A. M., Wang, J. W. & Axel, R. Spatial representation of the glomerular map in the *Drosophila* protocerebrum. *Cell* **109**, 229–241 (2002).
44. Keppel, G. & Wickens, T. D. *Design and Analysis: A Researcher's Handbook* 4th edn (Prentice Hall, 2004).

©2009 Macmillan Publishers Limited. All rights reserved

ARTICLES

Membrane scission by the ESCRT-III complex

Thomas Wollert[1]*, Christian Wunder[2]*, Jennifer Lippincott-Schwartz[2] & James H. Hurley[1]

The endosomal sorting complex required for transport (ESCRT) system is essential for multivesicular body biogenesis, in which cargo sorting is coupled to the invagination and scission of intralumenal vesicles. The ESCRTs are also needed for budding of enveloped viruses including human immunodeficiency virus 1, and for membrane abscission in cytokinesis. In *Saccharomyces cerevisiae*, ESCRT-III consists of Vps20, Snf7, Vps24 and Vps2 (also known as Did4), which assemble in that order and require the ATPase Vps4 for their disassembly. In this study, the ESCRT-III-dependent budding and scission of intralumenal vesicles into giant unilamellar vesicles was reconstituted and visualized by fluorescence microscopy. Here we show that three subunits of ESCRT-III, Vps20, Snf7 and Vps24, are sufficient to detach intralumenal vesicles. Vps2, the ESCRT-III subunit responsible for recruiting Vps4, and the ATPase activity of Vps4 were required for ESCRT-III recycling and supported additional rounds of budding. The minimum set of ESCRT-III and Vps4 proteins capable of multiple cycles of vesicle detachment corresponds to the ancient set of ESCRT proteins conserved from archaea to animals.

The ESCRT machinery is essential for several fundamental cellular pathways[1–7]. The biogenesis of lysosomes involves the maturation of early endosomes into multivesicular bodies (MVBs)[1,2,5]. In this pathway, portions of the limiting membrane of the endosome invaginate and then detach into the lumen of the endosome, forming intralumenal vesicles (ILVs). The MVBs then fuse with the lysosome and the ILVs and their contents are degraded. Cell-surface receptors destined for downregulation and some lysosomal-resident enzymes are sorted into this pathway following their covalent ubiquitination. The ESCRTs are involved both in sorting of ubiquitinated cargo into ILVs and in the morphogenesis of the ILVs themselves[1,2,5,7]. The ESCRT complexes are also required for the release of nascent human immunodeficiency virus 1 (HIV-1) virions from the plasma membrane of human cells[8–10]. The ESCRTs are needed for cytokinesis in both animal cells[11,12] and a subset of archaea[13,14].

The detachment of ILVs from the endosomal limiting membrane involves the scission of the narrow membrane neck connecting the nascent vesicle with the limiting membrane (Supplementary Fig. 1). This process is topologically opposite to the mechanism of membrane cleavage by dynamin family proteins[15], which coat the outer surface of membrane tubules and cleave from the outside. Cleavage of the membrane neck connecting a nascent virion to the plasma membrane or connecting two daughter cells in cytokinesis also occurs from the membrane surface contiguous with the inside of the neck (Supplementary Fig. 1). The conserved and ubiquitous involvement of ESCRTs in the inwardly directed cleavage of membrane necks has led to the view that ESCRTs are likely to have membrane scission activity. So far the absence of an ESCRT scission assay reconstituted entirely from purified proteins and lipids has prevented a direct test of this hypothesis, and has more broadly impeded progress in the ESCRT field.

The ESCRT family consists of the ESCRT-0, -I, -II and -III complexes together with Vps4, Alix (Bro1 in yeast) and other associated proteins[4,6,7]. ESCRT-0, -I and -II are soluble, contain ubiquitin-binding domains, and are primarily involved in cargo sorting and in the recruitment and activation of ESCRT-III. The ESCRT-III subunits Vps20, Snf7, Vps24 and Vps2 are soluble monomers in cytosol[16], and the crystal structure of one of these subunits is known[17]. Activated ESCRT-III subunits assemble into a detergent-insoluble complex on the endosomal membrane[16,18,19]. Overexpression of two of the human orthologues of Snf7 (CHMP4A and CHMP4B) leads to plasma membrane evaginations that are coated on the inside and have dimensions comparable to those of membrane necks cleaved by the ESCRTs[20]. ESCRT-III disassembly requires hydrolysis of ATP by the AAA ATPase Vps4 (ref. 21). The amino-terminal MIT domain of Vps4 binds to a subset of ESCRT-III proteins by means of MIM1 (refs 22, 23) and MIM2 (ref. 24) motifs in their carboxy termini. ESCRT-III subunits can form filaments[25] and helical tubes *in vitro*[26] that can be disassembled by Vps4. These observations indicate that the ESCRT-III−Vps4 module is a good candidate for a membrane scission machine. Recently, fluorescence microscopy of bilayer giant unilamellar vesicles (GUVs) has become established as a tool for the reconstitution of reactions in membrane trafficking[27–30]. We set out to initiate an *in vitro* assay for the ESCRT machinery and to determine whether the ESCRTs have an intrinsic membrane scission capability using purified ESCRT-III subunits and Vps4 in a GUV-based assay.

ESCRT-III and Vps4 recruitment *in vitro*

Purified ESCRT-III proteins aggregate when overexpressed in bacteria or when purified and concentrated; this has posed an obstacle to biochemical analysis. To avoid off-pathway effects due to aggregation, care was taken to preserve all ESCRT-III subunits in their monomeric state. ESCRT-III subunits were expressed as cleavable maltose-binding protein (MBP) fusions[26], and maintained at low micromolar concentrations following cleavage. The absence of aggregation was verified for each ESCRT-III protein by size-exclusion chromatography (SEC; Supplementary Fig. 2).

In the yeast MVB pathway, Vps20 is recruited and activated by ESCRT-II (refs 16, 31). Vps20 and other ESCRT-III proteins are maintained in an inactive, soluble state by autoinhibition by their C termini[18,19]. C-terminal truncations of the autoinhibitory domains

[1]Laboratory of Molecular Biology, National Institute of Diabetes and Digestive and Kidney Diseases, and [2]Cell Biology and Metabolism Branch, National Institute of Child Health and Human Development, National Institutes of Health, US Department of Health and Human Services, Bethesda, Maryland 20892, USA.
*These authors contributed equally to this work.

©2009 Macmillan Publishers Limited. All rights reserved

render ESCRT-III subunits constitutively active *in vivo*. To simplify the reconstitution of the system by bypassing the need for ESCRT-II while avoiding spontaneous aggregation, we used the partial C-terminal deletion construct Vps20ΔC. This construct was monomeric, as judged by SEC and light scattering (Supplementary Fig. 2).

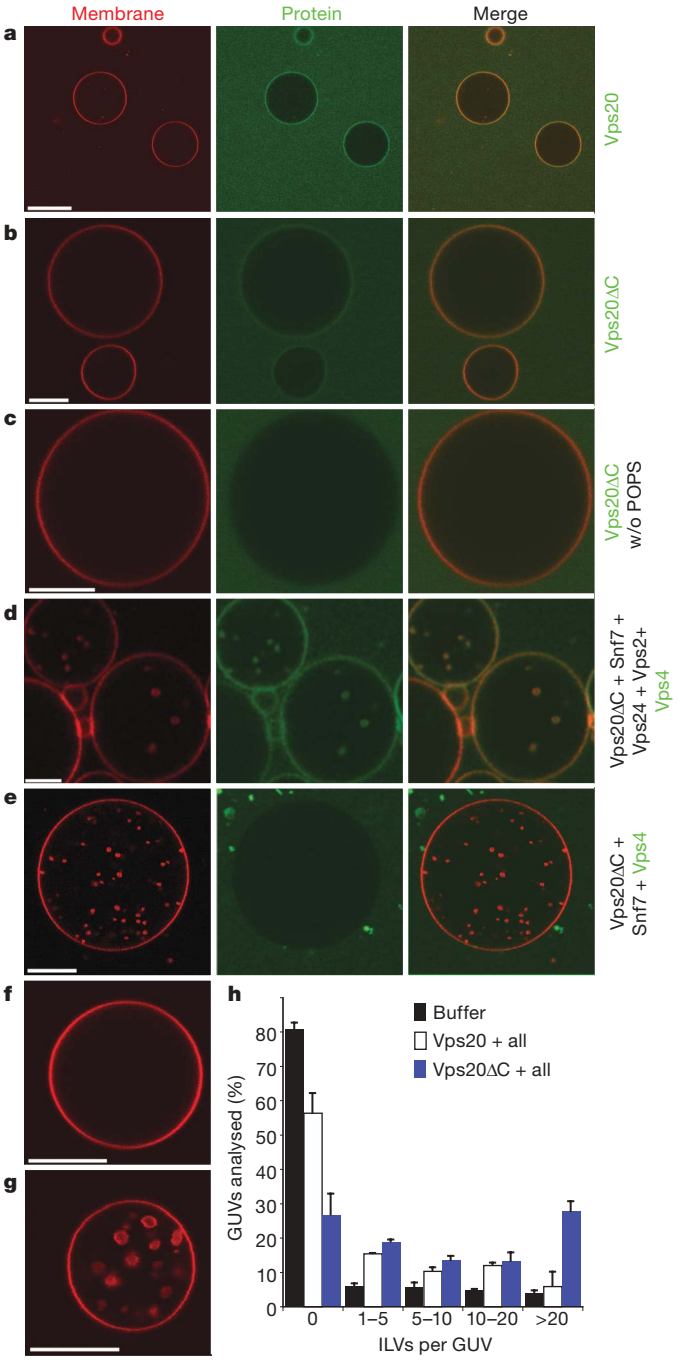


Figure 1 | ESCRT-III and Vps4 bind to GUV membranes. The GUV membrane is stained red with rhodamine-PE, and selected proteins are labelled green with Alexa 488. **a**, **b**, Vps20 (**a**) and Vps20ΔC (**b**) are recruited to membranes containing 10% POPS and 3% PtdIns(3)P. **c**, Vps20ΔC is not recruited to membranes lacking (w/o) POPS. **d**, Vps4 colocalizes with GUVs in the presence of Vps20ΔC, Snf7, Vps2 and Vps24. **e**, Vps4 is not recruited to GUVs if Vps2 and Vps24 are omitted. **f**, Five rounds of buffer addition to GUVs leads to no ILV formation. **g**, Vps20ΔC, Snf7, Vps24, Vps2 and Vps4 generate ILVs. **h**, ILVs were counted in a total 100 GUVs for each condition shown. Error bars were calculated using two independent experiments and calculating the difference between them. 'All' refers to Snf7, Vps24, Vps2 and Vps4. Scale bars, 10 μm.

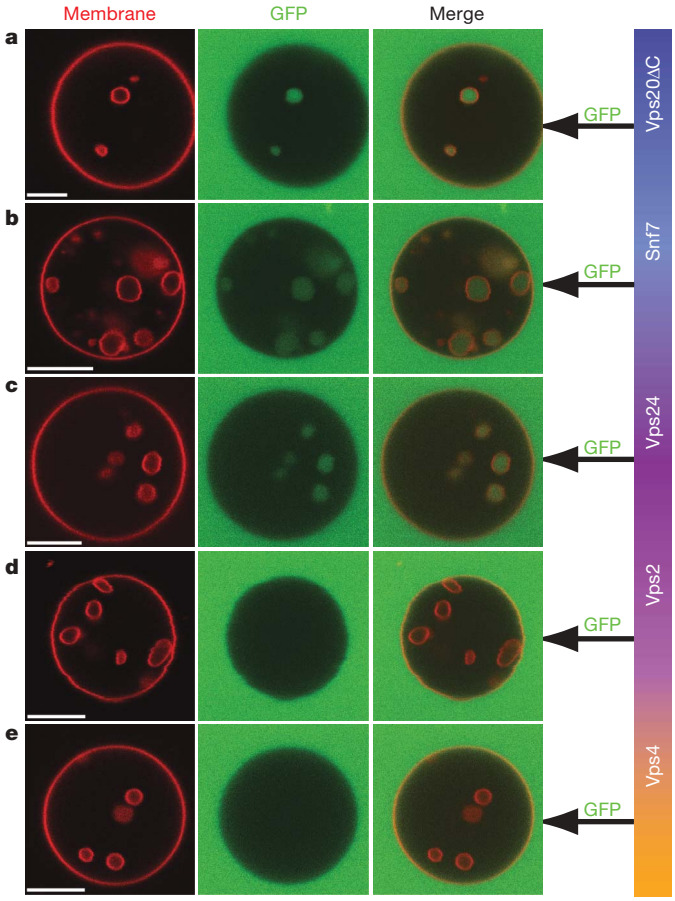


Figure 2 | Uptake of the bulk phase and vesicle scission by ESCRT-III. GUVs (red) were incubated with Vps20ΔC, Snf7, Vps24, Vps2 and Vps4 in each experiment shown. Each component was added sequentially and incubated for 5 min before the addition of the next component. To determine whether ILVs are connected with the bulk phase, the soluble marker GFP (green) was added immediately after Vps20ΔC (**a**), Snf7 (**b**), Vps24 (**c**), Vps2 (**d**) or Vps4 (**e**). No ATP or other nucleotide was present at any time. These data show that, until Vps2 was added, the ILVs were connected with the bulk phase. Scale bars, 10 μm.

GUVs with an endosome-like membrane composition[32,33] were generated by electroformation[34] from 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC):1-palmitoyl-2-oleoyl-sn-glycero-3-phospho-l-serine (POPS):phosphatidylinositol 3-phosphate (PtdIns(3)P):cholesterol at a molar ratio of 62:10:3:25. We verified that fluorescently labelled Vps20 and Vps20ΔC bound to these GUVs (Fig. 1a, b). Vps20 did not bind efficiently when the acidic lipid POPS was omitted from the mixture (Fig. 1c), consistent with the expectation that Vps20 binds to negatively charged membranes through a large electropositive surface[17]. Yeast ESCRT-III assembles in the following order: Vps20, Snf7, Vps24, Vps2 (ref. 35). Vps2 is the only one of these subunits to contain the Vps4-binding MIM1 motif[22,23]. The Vps20ΔC construct used in this study lacked the Vps4-binding MIM2 motif[24]. Vps4 is not recruited to membranes in the absence of ESCRT-III. Unlabelled ESCRT-IIIs were added sequentially and the recruitment of labelled Vps4 was monitored. Vps4 was recruited to the limiting membrane and internal structures in GUVs when all four ESCRT-III subunits were added beforehand (Fig. 1d). Vps4 was recruited neither to the limiting membrane nor to internal structures if Vps24 and Vps2 were omitted (Fig. 1e). The ESCRT-III subunits, Vps4 and the GUV model membrane thus support ordered ESCRT-III and Vps4 recruitment.

ESCRT-III promotes ILV formation

ESCRT subunits were added at submicromolar concentrations in the appropriate order, followed by Vps4. When the electroformed GUVs

©2009 Macmillan Publishers Limited. All rights reserved

were shifted to the experimental buffer, a small amount of membrane invagination was observed. Therefore, all results were compared to controls that had undergone identical manipulations and incubations using buffer blanks (Fig. 1f). When the reaction was initiated with Vps20ΔC, a massive increase in the number of ILVs was observed over baseline (Fig. 1g, h). Full-length Vps20 was less active, consistent with its autoinhibition (Fig. 1h). Vps20ΔC was used for all subsequent experiments owing to its greater activity. Labelled Vps4 was internalized into ILVs when all ESCRT-III subunits were present (Fig. 1d). Much or all of this pool of Vps4, and the unlabelled ESCRT-III associated with it, may be a passenger on the ILVs. In the absence of cargo, viral Gag or other material on the membrane, it would not be surprising for a portion of the ESCRT-III and Vps4 to associate non-productively with naked lipids. These data show that ESCRT-III proteins are able to markedly perturb membrane structure.

ESCRT-III carries out membrane scission

The diffusion of soluble green fluorescent protein (GFP) was used to assess permeability between bulk solution and putative ILVs. In each experiment, all of the ESCRT-III subunits and Vps4 were added, whereas the order of addition of GFP was varied (Fig. 2). When GFP was added with Vps20ΔC, Snf7 or Vps24, all of the ILVs imaged were green (Fig. 2a–c). Therefore, until Vps24 was added, all of the internal structures were still connected to the exterior through the bulk phase. When GFP was added with Vps2 (Fig. 2d), no green structures were observed. Thus, at this stage, all of the ILVs were detached from the limiting membrane. No further change in the nature of the ILVs was seen when Vps4 was added in the absence of ATP (Fig. 2e).

ILV detachment was corroborated by a three-dimensional reconstruction of a GUV in which GFP had been added at the Snf7 stage (Fig. 3a and Supplementary Movie 1). The ILVs in this GUV were green (Fig. 3b), indicating that they were formed in the course of the experiment. Another reconstruction was performed for a GUV treated with only Vps20, Snf7 and Vps24 (data not shown), showing complete detachment of ILVs. The ILVs within a GUV were highly mobile (Supplementary Movie 2), consistent with their complete detachment.

The bulk phase uptake experiment was repeated while omitting the subunits one at a time. GFP was added after Vps24 in these experiments. Omission of Vps20ΔC led to a sharp decrease in ILV formation (Fig. 4a), whereas omission of Snf7 almost completely abrogated ILV formation (Fig. 4b). Omission of Vps24, with GFP added in its place, led to modestly decreased ILV formation (Fig. 4c). A combination of detached and non-detached vesicles was observed in the absence of Vps24 (Fig. 4c). When Vps2 or Vps4 was omitted (Fig. 4d, e), the GUVs were indistinguishable from the situation with

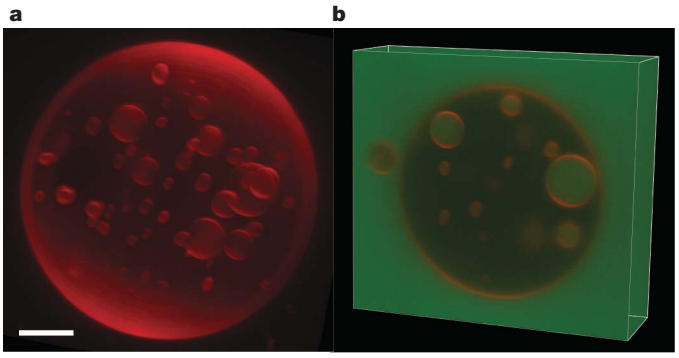


Figure 3 | Three-dimensional reconstruction of an ESCRT-III-treated GUV. **a**, Three-dimensional reconstruction of a GUV that was incubated with Vps20ΔC, Snf7, Vps24, Vps2 and Vps4, with GFP added together with Snf7. **b**, A *Z*-stack of the same GUV shows that the ILVs are filled with GFP, indicating that they were detached subsequent to the addition of Snf7. Scale bars, 5 μm.

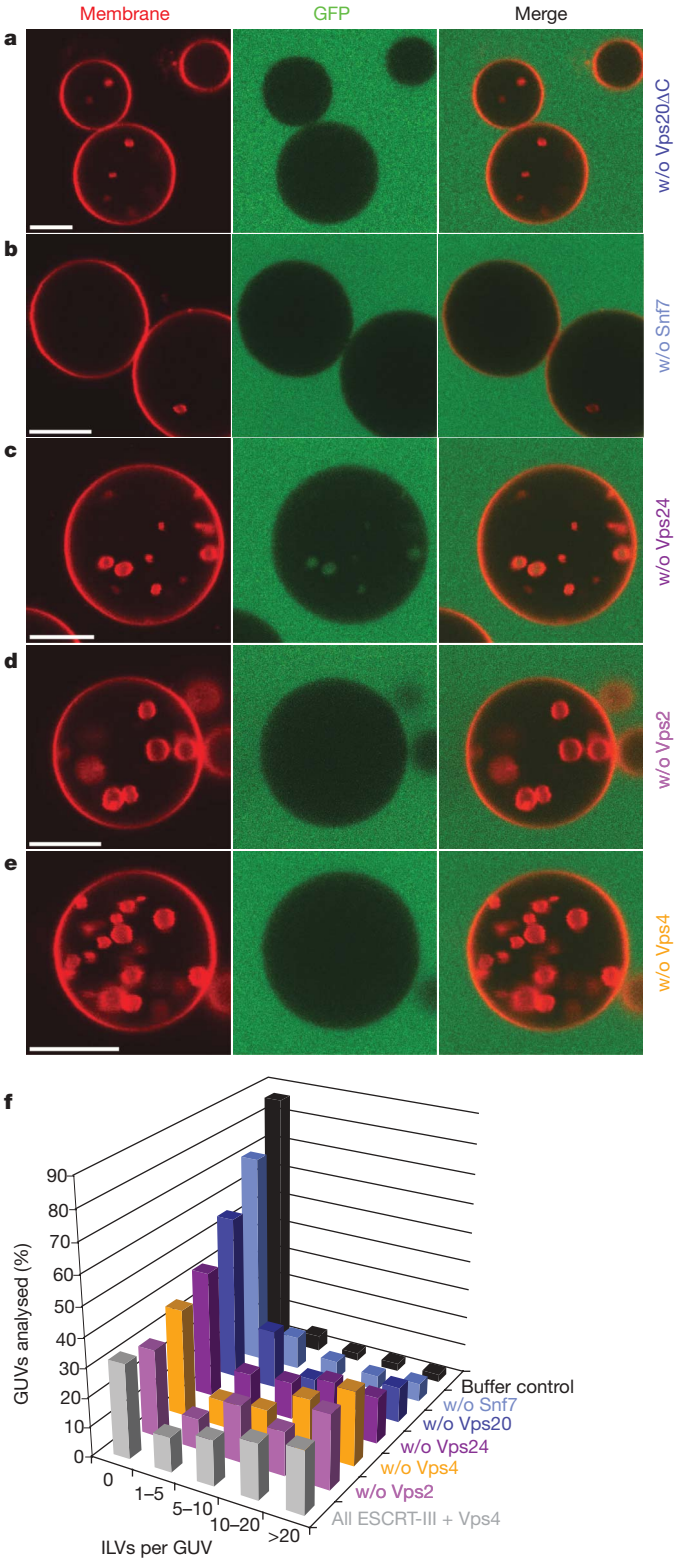


Figure 4 | Contributions of individual ESCRT-III subunits to ILV formation. GUVs were incubated with Vps20ΔC, Snf7, Vps24, Vps2 and Vps4, with one component omitted at a time. Omission of Vps20ΔC (**a**) or Snf7 (**b**) reduces ILV formation markedly. Omitting Vps24 (**c**) leads to delayed detachment of ILVs whereas omission of Vps2 (**d**) and Vps4 (**e**) has no significant effect. **f**, Summary of the omission experiment ($n = 100$). Each colour corresponds to a single experiment in which the indicated component was omitted; in the buffer control, only buffer was added to GUVs (black). Grey bars show the positive control, in which the reaction was initiated in the presence of Vps20ΔC, Snf7, Vps24, Vps2 and Vps4. Scale bars, 10 μm.

©2009 Macmillan Publishers Limited. All rights reserved

all components (Fig. 1g). We found that efficient ILV formation and detachment require Vps20ΔC, Snf7 and Vps24 (Fig. 4f). When PtdIns(3)P was replaced by another endosomal phosphoinositide, phosphatidylinositol-3,5-bisphosphate (PtdIns(3,5)P_2), the reaction proceeded just as with PtdIns(3)P (Supplementary Fig. 3a, c). When phosphoinositides were omitted but the mole fraction of the mono-anionic lipid POPS was increased to maintain the same net negative charge, ILV formation was reduced by about half (Supplementary Fig. 3b, c). Thus, phosphoinositides promote the reaction but are not strictly required.

Vps4 recycles ESCRT-III after scission

The bulk phase uptake experiment was modified such that ATP and GFP were added after Vps4 (Fig. 5a). The resulting images showed that the GUVs contained a mixture of GFP-filled and empty vesicles. When ATP and GFP were added, but Vps4 omitted, no GFP-filled vesicles were observed (Fig. 5b), as seen in the presence of Vps4 and absence of ATP (Fig. 2e). Thus, the addition of ATP to the ESCRT-III−Vps4 mixture after the termination of one round of ILV formation leads to a second wave of ILV formation. Omission of Vps2 completely inhibited the second wave of ILV formation (Fig. 5c, d), even though Vps2 is not essential for the initial round of formation. This is consistent with expectation, because Vps2 is the only ESCRT-III component in the model system that contains a Vps4-binding MIM motif[22,23]. Specific ESCRT-III−Vps4 interactions are required for recycling, recapitulating the biological function of Vps4 (ref. 16). Vps4 is functionally competent in the *in vitro* system, and the lack of a Vps4 effect on ILV detachment was not due to a lack of Vps4 activity. These observations show that the ATP hydrolysis step is not directly involved in vesicle detachment, but that it is indirectly involved by virtue of recycling ESCRT-III for multiple rounds of vesicle detachment.

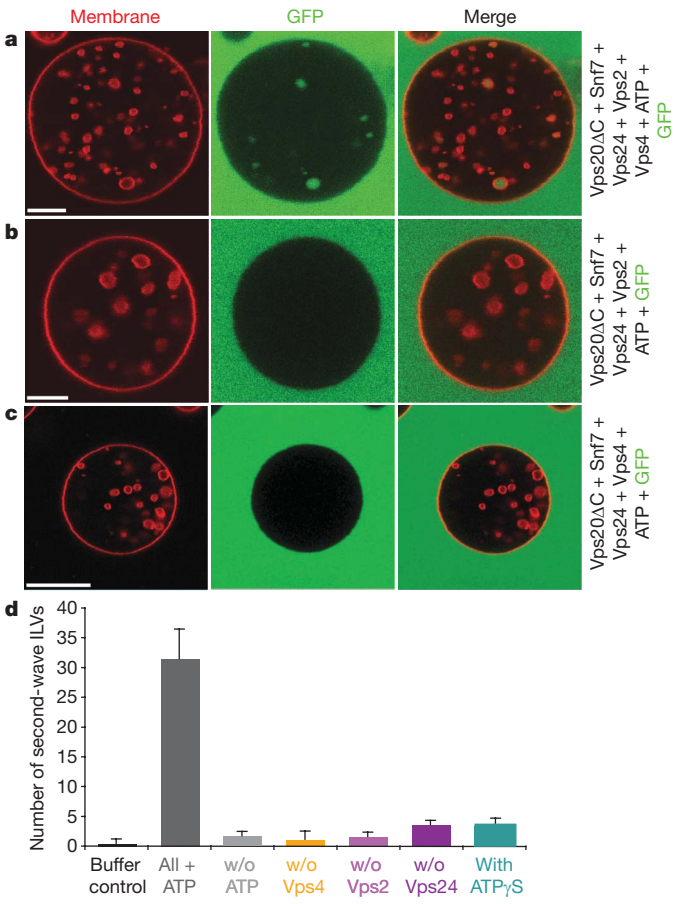


Figure 5 | Vps4 and ATP induce a second cycle of ILV formation. a, GUVs (red) were incubated with Vps20ΔC, Snf7, Vps24, Vps2 and Vps4. GFP (green) was added after Vps4. In the absence of Vps4 (**b**) or Vps2 (**c**), no GFP-filled ILVs are formed following ATP addition. Scale bars, 10 μm. **d**, Quantification of the ATP-induced second wave of ILV formation. Buffer control: only buffer and ATP were added to GUVs (black bar). Grey bars: GUVs with all protein components (Vps20ΔC, Snf7, Vps24, Vps2 and Vps4) in the presence (dark grey) and absence (light grey) of ATP. Coloured bars: the indicated subunit was omitted from the reaction in the presence of all other components and ATP. Two sets of 15 GUVs were analysed for each reaction and the difference between the two sets was used to calculate the error bars.

Discussion

We have reconstituted ESCRT-dependent membrane scission using a minimal set of protein components and a simple model for the eukaryotic endosomal membrane. The proteins included correspond to the minimal set conserved between eukaryotes and archaea, ESCRT-III and Vps4 (refs 13, 14). The model system recapitulates the ordered recruitment of ESCRT-III and Vps4 established in yeast. Snf7 has the highest stoichiometry *in vivo* relative to other subunits[35], and is the most important subunit in cargo sequestration[35]. The *in vitro* studies show that Snf7 is the most potent of the subunits in membrane scission. The behaviour of the system is consistent with known biological and biochemical data.

The first major insight is that ESCRT-III proteins alone have the intrinsic ability to drive the detachment of ILVs. ESCRT-III is capable of cleaving membrane necks with the appropriate inward topology in this minimal system. It is remarkable that the system does not require any of the upstream ESCRT complexes (-0, -I or -II), cargo, unusual lipids such as lysobisphosphatidic acid[27] or manipulation of membrane tension. Most of these factors are still likely to modulate the reaction, even if they are not essential. For example, the ILVs seen in this study are more heterogeneous and far larger than those seen physiologically[36]. We previously suggested that ESCRT-I and perhaps other upstream ESCRTs might regulate ILV size[37]. In a cellular setting the bud neck is scaffolded by upstream ESCRTs and cargo (MVB pathway), viral Gag assembly (HIV-1 budding) or the extension of the central spindle coupled to membrane addition (cytokinesis). In the *in vitro* assay the ESCRT-III proteins scaffold the formation of their own membrane neck. *In vitro* scission may be acting at both the intravesicular and extravesicular end of the neck, but the recycling step requires action from the extravesicular side. Substantial amounts of labelled ESCRT-III and Vps4 are on ILVs, indicating that some of the cleavage is occurring intravesicularly. However, the ability of Vps4 to recycle a portion of the ESCRT-III demonstrates that at least some *in vitro* scission is driven from the extravesicular end. Under physiological conditions, the presence of cargo, Gag, and so on, is likely to inhibit scission from the intravesicular side of the neck. The efficiency of the neck closure and the predominance of extravesicular cleavage would be higher in the physiological setting of a preformed neck. None of this alters the major conclusion that ESCRT-III alone can detach ILVs from the limiting membrane.

The second major insight pertains to the role of ATP hydrolysis by Vps4, the only thermodynamic driving force for the reaction. It has been suggested that ATP hydrolysis by Vps4, while anchored tightly to the membrane by ESCRT-III, could power a dynamin-like constriction of the neck leading to cleavage[26]. Although Vps4 can disassemble ESCRT-III *in vitro*[25,26], the relationship between disassembly and vesicle scission was unknown. Here we find that Vps4 acts after the membrane scission step to recycle ESCRT-III proteins. This has a precedent in SNARE-dependent vesicle fusion. SNAREs drive vesicle fusion by forming tight complexes that bring opposing membranes into proximity[38]. SNARE complexes do not dissociate spontaneously, but require the hydrolysis of ATP by the ATPase NSF (*N*-ethylmaleimide-sensitive factor) after membrane fusion[38]. A recycling role for Vps4 is consistent with all available data, and fits well with the SNARE−NSF precedent.

Snf7 is the most crucial of the ESCRT-III subunits for membrane scission. Snf7 can form either spirals or extended tubes when overexpressed[20]. The tubes suggest a mechanism for the formation of

©2009 Macmillan Publishers Limited. All rights reserved

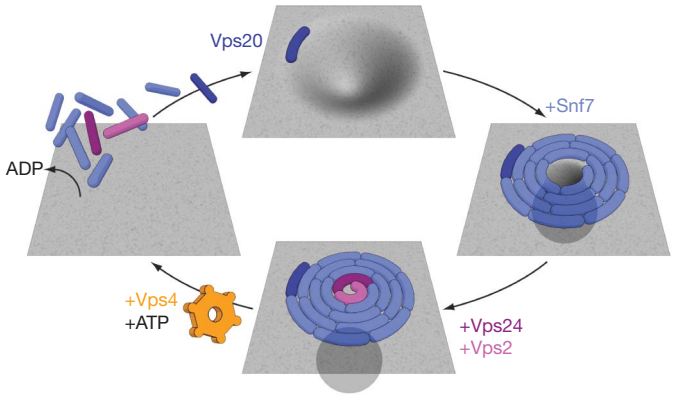


Figure 6 | A simple hypothesis for the mechanism of the ESCRT-III–Vps4 membrane scission cycle. The spiral structure is inspired by the observations of ref. 20. ESCRT-III sequesters cargo in MVBs, with Snf7 having the most important role of the four subunits[35]. Any cargo present in the neck near the growing array would be forced out of the neck and further into the bud at this stage, leading to its sequestration. Completion of the spiral would bring the membrane lipids of the neck into critical proximity, allowing fusion. The fused neck would then resolve into the detached vesicle on the interior and the ESCRT-III-coated limiting membrane on the exterior (bottom panel).

invaginations of the GUV membrane, although we do not propose that the tubes are involved in physiological budding. In MVB biogenesis the inward bud is filled with cargo, and in HIV-1 budding it is filled with the Gag protein. In the *in vitro* system, the ability of ESCRT-III components to form inward buds independently of cargo, Gag or other factors has fortuitously yielded membrane necks for study as scission substrates. The spirals suggest an attractive mechanism for membrane scission in both *in vitro* and physiological settings (Fig. 6). In this model, one surface of Snf7 is tightly anchored to the membrane. Other Snf7 surfaces form an inwardly growing spiral. The growing spiral could progressively squeeze the membrane neck until the critical proximity for membrane fusion is reached (Fig. 6 and Supplementary Movie 3).

In conclusion, ESCRT-III has the intrinsic ability to drive the scission of membrane necks. Vps4 recycles ESCRT-III but does not directly participate in neck closure. Further elaborations of this ESCRT reconstitution should help complete the mechanistic analysis of the pathway.

METHODS SUMMARY

Preparation of giant unilamellar vesicles. A thin film of lipid mixtures containing POPC (62 mol%), POPS (10 mol%), cholesterol (25 mol%), PtdIns(3)P (3 mol%) and 1,2-dipalmitoyl-sn-glycero-3-phosphoethanolamine-*N*-(lissamine rhodamine B sulphonyl) (rhodamine-PE) (0.1 mol%) was applied to indium-tin-oxide-covered glass slides and placed into a custom-made Teflon chamber. The glass slides were separated by a 2 mm thin glucose solution (600 mM), and an electric AC-field (1 V, 10 Hz) was applied for 4 h at 60 °C following the previously described GUV electro-formation protocol[34]. The GUVs were collected after cooling to room temperature (25 °C) and used immediately.

ESCRT reactions. A 200 µl observation chamber (Lab-Tek chambered #1.0 Borosilicate) was coated using 1 mg ml^{-1} BSA to avoid distortion of GUVs on contact with the bottom of the chamber and rinsed with buffer (50 mM Tris HCl, pH 7.4, 300 mM NaCl). One-hundred microlitres of GUVs was mixed with this buffer and incubated for 5 min. The buffer and all proteins added to GUVs matched the osmolarity of the GUV solution (~650 mosM due to evaporation of ~10% water during electro-formation at 60 °C). Vps20 (full length, residues 1–221) or Vps20ΔC (C-terminal truncation, residues 1–170) were added to the GUV–buffer mixture to yield a final concentration of 160 nM. The solution was stirred by a gentle rotation of the pipette tip, without aspiration. After 5 min of incubation at room temperature, Snf7 (final concentration of 600 nM), Vps24 (200 nM), Vps2 (200 nM) and Vps4 (200 nM) were added in that order, separated by 5-min incubation intervals. For ATP experiments, ATP and $MgCl_2$ (2 mM final concentration) were added to the GUV/protein mixture. To establish uptake of the bulk phase, the soluble marker GFP was added as indicated, yielding a final concentration of 1.65 µM. One buffer control was performed for each experiment. Stirring and incubation times were the same for the control and ESCRT-III additions, and GFP was added to the control and ESCRT-III experiments in an identical manner. For Vps20 or Vps20ΔC colocalization experiments a final concentration of 30 nM of Alexa-488-conjugated protein was used, but no other proteins were added. For Vps4 colocalization, the experiment was set up as described above, however, a mixture of unlabelled and labelled Vps4 (5:1) was added to a final total Vps4 concentration of 200 nM.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 7 January; accepted 30 January 2009.
Published online 22 February 2009.

1. Gruenberg, J. & Stenmark, H. The biogenesis of multivesicular endosomes. *Nature Rev. Mol. Cell Biol.* **5,** 317–323 (2004).
2. Russell, M. R. G., Nickerson, D. P. & Odorizzi, G. Molecular mechanisms of late endosome morphology, identity and sorting. *Curr. Opin. Cell Biol.* **18,** 422–428 (2006).
3. Slagsvold, T., Pattni, K., Malerod, L. & Stenmark, H. Endosomal and non-endosomal functions of ESCRT proteins. *Trends Cell Biol.* **16,** 317–326 (2006).
4. Saksena, S., Sun, J., Chu, T. & Emr, S. D. ESCRTing proteins in the endocytic pathway. *Trends Biochem. Sci.* **32,** 561–573 (2007).
5. Piper, R. C. & Katzmann, D. J. Biogenesis and function of multivesicular bodies. *Annu. Rev. Cell Dev. Biol.* **23,** 519–547 (2007).
6. Williams, R. L. & Urbe, S. The emerging shape of the ESCRT machinery. *Nature Rev. Mol. Cell Biol.* **8,** 355–368 (2007).
7. Hurley, J. H. ESCRT Complexes and the biogenesis of multivesicular bodies. *Curr. Opin. Cell Biol.* **20,** 4–11 (2008).
8. Morita, E. & Sundquist, W. I. Retrovirus budding. *Annu. Rev. Cell Dev. Biol.* **20,** 395–425 (2004).
9. Bieniasz, P. D. Late budding domains and host proteins in enveloped virus release. *Virology* **344,** 55–63 (2006).
10. Fujii, K., Hurley, J. H. & Freed, E. O. Beyond Tsg101: the role of Alix in 'ESCRTing' HIV-1. *Nature Rev. Microbiol.* **5,** 912–916 (2007).
11. Carlton, J. G. & Martin-Serrano, J. Parallels between cytokinesis and retroviral budding: a role for the ESCRT machinery. *Science* **316,** 1908–1912 (2007).
12. Morita, E. *et al.* Human ESCRT and ALIX proteins interact with proteins of the midbody and function in cytokinesis. *EMBO J.* **26,** 4215–4227 (2007).
13. Samson, R. Y., Obita, T., Freund, S. M., Williams, R. L. & Bell, S. D. A role for the ESCRT system in cell division in Archaea. *Science* **322,** 1710–1713 (2008).
14. Lindas, A. C., Karlsson, E. A., Lindgren, M. T., Ettema, T. J. G. & Bernander, R. A unique cell division machinery in the Archaea. *Proc. Natl Acad. Sci. USA* **105,** 18942–18946 (2008).
15. Hinshaw, J. E. Dynamin and its role in membrane fission. *Annu. Rev. Cell Dev. Biol.* **16,** 483–519 (2000).
16. Babst, M., Katzmann, D. J., Estepa-Sabal, E. J., Meerloo, T. & Emr, S. D. ESCRT-III: An endosome-associated heterooligomeric protein complex required for MVB sorting. *Dev. Cell* **3,** 271–282 (2002).
17. Muziol, T. *et al.* Structural basis for budding by the ESCRT-III factor CHMP3. *Dev. Cell* **10,** 821–830 (2006).
18. Zamborlini, A. *et al.* Release of autoinhibition converts ESCRT-III components into potent inhibitors of HIV-1 budding. *Proc. Natl Acad. Sci. USA* **103,** 19140–19145 (2006).
19. Shim, S., Kimpler, L. A. & Hanson, P. I. Structure/function analysis of four core ESCRT-III proteins reveals common regulatory role for extreme C-terminal domain. *Traffic* **8,** 1068–1079 (2007).
20. Hanson, P. I., Roth, R., Lin, Y. & Heuser, J. E. Plasma membrane deformation by circular arrays of ESCRT-III protein filaments. *J. Cell Biol.* **180,** 389–402 (2008).
21. Babst, M., Wendland, B., Estepa, E. J. & Emr, S. D. The Vps4p AAA ATPase regulates membrane association of a Vps protein complex required for normal endosome function. *EMBO J.* **17,** 2982–2993 (1998).
22. Obita, T. *et al.* Structural basis for selective recognition of ESCRT-III by the AAA ATPase Vps4. *Nature* **449,** 735–739 (2007).
23. Stuchell-Brereton, M. *et al.* ESCRT-III recognition by VPS4 ATPases. *Nature* **449,** 740–744 (2007).
24. Kieffer, C. *et al.* Two distinct modes of ESCRT-III recognition are required for VPS4 functions in lysosomal protein targeting and HIV-1 budding. *Dev. Cell* **15,** 62–73 (2008).
25. Ghazi-Tabatabai, S. *et al.* Structure and disassembly of filaments formed by the ESCRT-III subunit Vps24. *Structure* **16,** 1345–1356 (2008).
26. Lata, S. *et al.* Helical structures of ESCRT-III are disassembled by VPS4. *Science* **321,** 1354–1357 (2008).
27. Matsuo, H. *et al.* Role of LBPA and Alix in multivesicular liposome formation and endosome organization. *Science* **303,** 531–534 (2004).
28. Trajkovic, K. *et al.* Ceramide triggers budding of exosome vesicles into multivesicular endosomes. *Science* **319,** 1244–1247 (2008).
29. Sens, P., Johannes, L. & Bassereau, P. Biophysical approaches to protein-induced membrane deformations in trafficking. *Curr. Opin. Cell Biol.* **20,** 476–482 (2008).
30. Romer, W. *et al.* Shiga toxin induces tubular membrane invaginations for its uptake into cells. *Nature* **450,** 670–675 (2007).

©2009 Macmillan Publishers Limited. All rights reserved

31. Teo, H., Perisic, O., Gonzalez, B. & Williams, R. L. ESCRT-II, an endosome-associated complex required for protein sorting: crystal structure and interactions with ESCRT-III and membranes. *Dev. Cell* **7,** 559–569 (2004).
32. Evans, W. H. & Hardison, W. G. Phospholipid, cholesterol, polypeptide and glycoprotein composition of hepatic endosome subfractions. *Biochem. J.* **232,** 33–36 (1985).
33. Kobayashi, T. *et al.* Separation and characterization of late endosomal membrane domains. *J. Biol. Chem.* **277,** 32157–32164 (2002).
34. Angelova, M. I. & Dimitrov, D. S. Liposome electroformation. *Faraday Discuss. Chem. Soc.* **81,** 303–311 (1986).
35. Teis, D., Saksena, S. & Emr, S. D. Ordered assembly of the ESCRT-III complex on endosomes is required to sequester cargo during MVB formation. *Dev. Cell* **15,** 578–589 (2008).
36. Nickerson, D. P., West, M. & Odorizzi, G. Did2 coordinates Vps4-mediated dissociation of ESCRT-III from endosomes. *J. Cell Biol.* **175,** 715–720 (2006).
37. Kostelansky, M. S. *et al.* Molecular architecture and functional model of the complete yeast ESCRT-I heterotetramer. *Cell* **129,** 485–498 (2007).
38. Wickner, W. & Schekman, R. Membrane fusion. *Nature Struct. Mol. Biol.* **15,** 658–664 (2008).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank V. Schram and the NICHD imaging core facility for use of the LSM5 LiveDuo microscope, D. Yang and G. Patterson for providing purified Vps4 and GFP, respectively, B. Beach for assistance with subcloning, E. Tyler for generating Fig. 6 and Supplementary Movie 3, Y. Im for assistance with Supplementary Fig. 1, C. Biertümpfel and W. Yang for use of their light-scattering instrument, and members of the Hurley and Lippincott–Schwartz laboratories for discussions. This research was supported by the Intramural Program of the National Institutes of Health, NICHD (J.L.-S.), NIDDK (J.H.H.) and IATAP (J.H.H. and J.L.-S.). T.W. is an EMBO long-term fellow.

Author Contributions T.W. and C.W. prepared GUVs and carried out confocal microscopy; T.W. purified and labelled proteins; C.W. carried out the three-dimensional reconstruction; T.W., C.W., J.L.-S. and J.H.H. analysed data; J.H.H. designed the study and wrote the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to J.H.H. (hurley@helix.nih.gov).

©2009 Macmillan Publishers Limited. All rights reserved

METHODS

Cloning, expression and purification of proteins. The ESCRT-III subunits Vps20, Snf7, Vps24 and Vps2 were subcloned into a modified pHis2 expression vector, containing the *Escherichia coli* MBP gene in between a 5′ His_6-tag and a 3′ TEV protease cleavage site. The C-terminally truncated version of Vps20 (Vps20ΔC, residues 1−170) was generated by PCR using the Vps20 construct above as a template and was subcloned as for Vps20. All constructs were transformed into Rosetta pLys cells (Stratagene). N-terminal His_6−MBP fusion proteins were expressed at 30 °C for 3 h in LB (Luria Bertani) medium supplemented with ampicillin (100 μg ml^{-1} final concentration) after induction with 0.5 mM IPTG at an optical density ($D_{600\,nm}$) of 0.8. The proteins were affinity purified using Ni-NTA resin (Qiagen) and eluted with 100 mM imidazole. Elution fractions were cleaved using TEV protease and further purified by SEC (Superdex 200 column) and a second round of chromatography on Ni-NTA resin. Pooled fractions of Snf7 were not concentrated to avoid aggregation of the monomeric protein. The final Snf7 protein concentration was usually between 4 and 8 μM. All other proteins were concentrated until ~50 μM. Purified proteins were immediately flash-frozen in liquid nitrogen and stored at −80 °C until use. The monomeric state of all proteins was verified using analytical SEC before usage (Superdex 200 10/30 column; GE Healthcare).

Full-length yeast Vps4 was purified from *E. coli* as a glutathione *S*-transferase (GST) fusion protein. After affinity-purification using glutathione resin and cleavage with TEV protease, the protein was subjected to SEC, and remnants of GST as well as TEV protease were removed by a second round of chromatography using Ni-NTA and glutathione resin. The purified protein was concentrated to 50 μM and stored at −80 °C.

Light-scattering analysis. Protein samples (5.5 mg ml^{-1} of Vps20 and 1.7 mg ml^{-1} of Vps20ΔC) were injected onto a TSK-GEL GMPWXL SEC (Tosoh) attached to an Agilent 1200 HPLC system (Agilent Technologies) using a 20 μl sample loop and a flow rate of 0.5 ml min^{-1}. The eluted peaks were analysed using a Wyatt DAWN HELEOS-II multi-angle light-scattering instrument and a Wyatt Optilab rEX differential refractometer. Data were evaluated using ASTRA 5.3.4 software using the Zimm model for static light scattering data fitting. Vps20 and Vps20ΔC were monodisperse and had experimental molecular weights of 29.4 and 24.1 kDa, respectively, as compared to their calculated molecular weights of 26.6 and 21.1 kDa.

Fluorescent labelling of proteins. To label Vps20 or Vps20ΔC, which contain no native Cys residues, the mutant N85C was generated using the QuickChange Mutagenesis Kit (Stratagene). The purified and concentrated protein solution was labelled by incubating Alexa Fluor 488 C5 maleimide (Molecular Probes) at a molar ratio of 1:1 with the protein at 37 °C for 1 h under a N_2 atmosphere. The labelling incorporation was evaluated by running samples on an SDS-PAGE gel and carrying out fluorescence scans of the gel with a Typhoon variable mode imager (GE-Healthcare) using 488 nm excitation and 495 nm detection wavelengths. Quantification of labelling efficiency was achieved by separating excess dye from labelled protein using two consecutive HiTrap desalting columns (GE-Healthcare). Molar protein and Alexa Fluor 488 concentrations of the purified protein were determined using absorption at 280 and 495 nm, respectively. Vps4 contains two cysteine residues at positions 317 and 376, which were used for the labelling reaction. The labelling and purification protocol was similar to that of the Vps20 constructs. Labelling efficiencies for all reactions exceeded 90%.

Lipids. The following lipids were purchased from Avanti Polar Lipids: POPS, POPC, cholesterol and 1,2-dipalmitoyl-sn-glycero-3-phosphoethanolamine-*N*-(lissamine rhodamine B sulphonyl) (rhodamine-PE). Phosphatidylinositol 3-phosphate and 3,5-bisphosphate (diC16) were purchased from Echelon. POPS, POPC, cholesterol and rhodamine-PE were dissolved in chloroform. Phosphoinositides were dissolved in 1:2:0.8 chloroform:methanol:water.

Confocal fluorescence microscopy. Images were taken in multi-tracking mode on a Zeiss LSM510 or LSM5 LiveDuo confocal microscope with a ×63 Plan Apochromat 1.4 NA objective and a 488/543 dichroic mirror at a resolution of 512 × 512 pixels. The GFP and Alexa 488 dyes were excited using the 488 nm line and the rhodamine was excited with a 543 nm HeNe laser. GFP and Alexa 488 emission was collected with a 505−530 nm bandpass filter. Rhodamine emission was collected with a 560 nm longpass filter. The pinholes for each channel were set for an approximate 1.5–2.5 μm optical slice. Laser power was 9 μW for the 543 nm channel and 24, 400 or 810 μW for GFP bulk phase experiments, Vps20 colocalization or Vps4 colocalization, respectively, in the 488 nm channel. Images were analysed using the LSM Examiner software.

Enumeration of ILV formation. Randomly chosen field of views were evaluated to reveal the number of ILVs per GUV. Reactions were performed as described above and all GUVs in the particular fields of view were scanned in the *z* direction. All ILV per GUV were counted, and for each experimental condition 100 GUVs were analysed. Results were summarized in histograms.

Three-dimensional fluorescence microscopy of GUVs. *Z*-stacks over time were taken with an LSM5 LiveDuo (Carl Zeiss MicroImaging Inc.) using ×63 1.4 NA Plan-Apochromate oil objective at 90 frames per second, 1,024 × 256 pixels and a piezo focusing motor for fast hyperfine sectioning (0.35 μm). Static and animated three-dimensional visualizations of the data stacks were produced with IMARIS 6.1.0 software package (Bitplane Inc.).

©2009 Macmillan Publishers Limited. All rights reserved

LETTERS

Sub-cycle switch-on of ultrastrong light–matter interaction

G. Günter[1], A. A. Anappara[1,2], J. Hees[1], A. Sell[1], G. Biasiol[3], L. Sorba[2,3], S. De Liberato[4,5], C. Ciuti[4], A. Tredicucci[2], A. Leitenstorfer[1] & R. Huber[1]

Controlling the way light interacts with material excitations is at the heart of cavity quantum electrodynamics (QED). In the strong-coupling regime, quantum emitters in a microresonator absorb and spontaneously re-emit a photon many times before dissipation becomes effective, giving rise to mixed light–matter eigenmodes[1–12]. Recent experiments[13] in semiconductor microcavities reached a new limit of ultrastrong coupling[14], where photon exchange occurs on timescales comparable to the oscillation period of light. In this limit, ultrafast modulation of the coupling strength has been suggested to lead to unconventional QED phenomena[14,15]. Although sophisticated light–matter coupling has been achieved in all three spatial dimensions, control in the fourth dimension, time, is little developed. Here we use a quantum-well waveguide structure to optically tune light–matter interaction from weak to ultrastrong and turn on maximum coupling within less than one cycle of light. In this regime, a class of extremely non-adiabatic phenomena becomes observable. In particular, we directly monitor how a coherent photon population converts to cavity polaritons during abrupt switching. This system forms a promising laboratory in which to study novel sub-cycle QED effects and represents an efficient room-temperature switching device operating at unprecedented speed.

Microcavities enclosing a material resonance provide an elegant way to tailor the strength of light–matter interaction, quantified by the vacuum Rabi frequency, Ω_R. Intuitively, Ω_R represents the rate at which a photon is exchanged between an optical transition and a cavity mode through spontaneous emission and absorption. The following three regimes are distinguished. (1) For weak coupling, the discrete density of photonic states modifies the radiative lifetime of the material excitation (Purcell effect[16]). (2) If Ω_R exceeds the dissipation rates of light and matter fields, two polariton branches emerge as new eigenstates featuring an energy splitting of $2\hbar\Omega_R$, where $\hbar$ denotes Planck's constant divided by 2π. This strong-coupling regime has been investigated in diverse systems, ranging from atoms[1,2] through excitons in semiconductor quantum wells[3] and quantum dots[5,9] to Cooper-pair boxes[6,10]. Exciting perspectives arise for quantum information processing[5,6,10], lasing without inversion[17] and polaritonic Bose–Einstein condensation[17,18]. (3) Most recently, the giant dipole moments of intersubband transitions in quantum wells[19] have allowed for ultrastrong light–matter coupling[13–15]. Here Ω_R is large enough to amount to a significant fraction of the transition frequency, ω_{12} (for example, $2\Omega_R/\omega_{12} = 0.2$ in ref. 13).

In this regime, anti-resonant terms of the interaction Hamiltonian, usually accounted for in strongly driven systems only, become relevant even in equilibrium. The ground state is predicted to be a squeezed vacuum containing a finite number of virtual photons[14]. As long as the quanta are confined inside the cavity, their unconventional nature remains hidden. Theory shows that non-adiabatic switching of Ω_R, introducing changes on a timescale shorter than the cycle of light, may release these virtual quanta in correlated pairs[14,15]. This example reveals a fundamental weakness of state-of-the-art light–matter control: up to now, non-adiabatic phenomena have remained an academic curiosity because there has been no laboratory capable of controlling light–matter interaction on a sub-cycle scale. Electronic tuning by means of an external gate voltage is clearly too slow[7]; theoretical proposals to modulate the dielectric constant of a cavity or the reflectivity of a mirror[20] have not been implemented for the required timescale either.

We introduce an all-optical scheme for femtosecond control of ultrastrongly coupled cavity polaritons—a test bed for targeting non-adiabatic light–matter interaction. The idea is sketched in Fig. 1. For maximum modulation depth, we start with an empty microcavity (Fig. 1a) and turn on coupling by direct injection of the emitter itself (Fig. 1b). Our sample contains 50 identical, undoped GaAs quantum wells separated by $Al_{0.33}Ga_{0.67}As$ barriers (Fig. 1c). The electronic wavefunctions are quantized along the growth direction forming subbands $|1\rangle$ and $|2\rangle$, which are unpopulated in thermal equilibrium. Radiative transitions between the subbands become possible only if electrons are injected into the first subband. The intersubband absorption line is centred at $\hbar\omega_{12} = 113$ meV (wavelength, $\lambda = 11\,\mu$m) with a dipole moment oriented along the growth direction. We embed the quantum wells in a planar waveguide for mid-infrared radiation. The effective thickness of the structure corresponds to $\lambda/2$ at an internal propagation angle of $\theta = 65^\circ$. Hence, photon modes with electric field components in the growth direction (TM polarization) couple resonantly to intersubband transitions as long as the subbands are populated. The vacuum Rabi frequency is known to scale with the electron sheet density, N_e, in level $|1\rangle$ as $\Omega_R \propto \sqrt{N_e}$ (refs 7, 14, 15). In previous work, N_e was provided by static doping or electronic injection, both of which are difficult to modulate with high bandwidths. We employ a 12-fs control pulse centred at a photon energy of 1.55 eV to excite electrons resonantly from the valence band into conduction subband $|1\rangle$ (Fig. 1c). Because the inverse frequency of the intersubband transition is 37 fs, we activate it within less than half a cycle of light.

The eigenmodes of this system are resonantly sampled by reflectance of a phase-stable mid-infrared pulse (blue curve, Fig. 1c) coupled through the prism-shaped substrate. To achieve sub-cycle resolution, the real-time oscillations of the probe field are mapped out by phase-matched electro-optic detection[21]. Such ultrabroadband terahertz optoelectronics (refs 22, 23 and references therein) has proved to be a powerful tool for investigating the field response of

[1]Department of Physics and Center for Applied Photonics, University of Konstanz, Universitätsstraße 10, 78464 Konstanz, Germany. [2]Laboratorio NEST, CNR-INFM and Scuola Normale Superiore, Piazza San Silvestro 12, I-56127 Pisa, Italy. [3]Laboratorio Nazionale TASC CNR-INFM, Area Science Park, I-34012 Trieste, Italy. [4]Laboratoire Matériaux et Phénomènes Quantiques, Université Paris Diderot-Paris 7, Case 7021, Bâtiment Condorcet, 75205 Paris, France. [5]Laboratoire Pierre Aigrain, Ecole Normale Supérieure, UMR 8551, 75005 Paris, France.

©2009 Macmillan Publishers Limited. All rights reserved

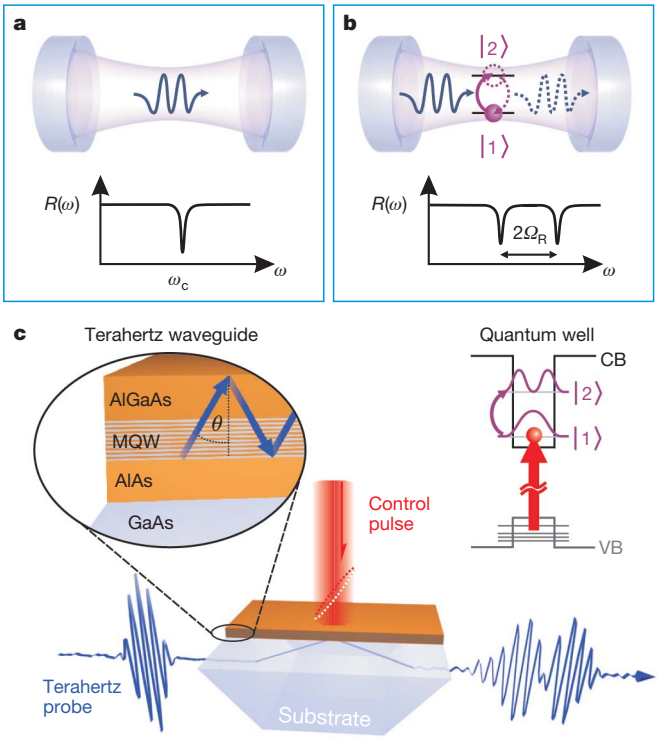


Figure 1 | Femtosecond control of ultrastrong light–matter coupling. **a**, A bare microcavity has minimal reflectance, $R(\omega)$, at the photon resonance, ω_c. **b**, After introduction of a resonant material excitation, cavity photons (blue) are coherently absorbed and re-emitted at rate Ω_R, giving rise to anticrossing cavity polaritons. **c**, A multiple quantum well structure (MQW) comprising 50 undoped GaAs wells (thickness, 9 nm) separated by $Al_{0.33}Ga_{0.67}As$ barriers (thickness, 30 nm) are embedded into a planar waveguide structure based on total internal reflection at the $Al_{0.33}Ga_{0.67}As$–air and AlAs–GaAs interfaces, respectively (magnified view). The quantum wells are positioned at the field antinode. The sketched band diagram (CB, conduction band; VB, valence band) shows how electronic transitions between subbands $|1\rangle$ and $|2\rangle$ (level spacing, $\hbar\omega_{12} = 113$ meV) are activated by near-infrared, 12-fs control pulses (photon energy, 1.55 eV; vertical red beam) populating level $|1\rangle$. Intersubband transitions may then resonantly couple to TM-polarized mid-infrared cavity photons propagating at $\theta = 65°$. Few-cycle TM-polarized multi-terahertz transients guided through the prism-shaped substrate are reflected from the waveguide to probe the ultrafast build-up of light–matter coupling in the system. The pulse front of the near-infrared control is tilted (dotted white circle in control beam) to match the geometry of the phase planes of the probe.

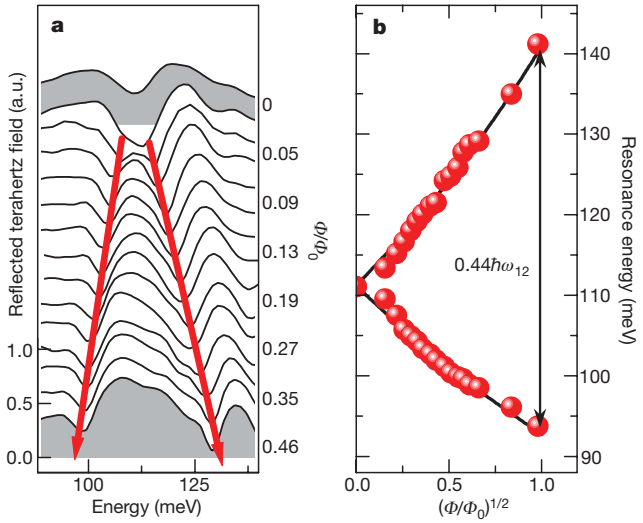


Figure 2 | Ultrawide optical tuning of light–matter interaction. **a**, Terahertz reflectance spectra measured at room temperature (293 K) for various fluences, Φ (vertically off-set), of the control pulse ($t_D = 20$ ps). Minima indicate eigenmodes of the system. For $\Phi = 0$, only the bare photon mode is observed, at $\hbar\omega_c = 113$ meV; both branches of the intersubband cavity polaritons are discernible for $\Phi \geq 0.05\Phi_0$ ($\Phi_0 = 0.1$ mJ cm^{-2}). a.u., arbitrary units. **b**, Asymmetric polariton splitting as a function of Φ: dots, experiment; solid lines, simulation including anti-resonant light–matter interaction. For $\Phi = \Phi_0$, the polariton branches are observed with a relative energy distance of $0.44\hbar\omega_{12}$ for a given angle, $\theta = 65°$. This value corresponds to $2\Omega_R = 0.18\omega_{12}$. We estimate the maximum electron density to be on the order of 2×10^{12} cm^{-2}, consistent with the static doping concentrations of ref. 13. The spectra are obtained by Fourier transformation of time-domain data shown in Supplementary Fig. 1.

extremely non-equilibrium semiconductor systems[24–26]. Eigenmodes of the cavity cause characteristic minima in the Fourier spectra of the reflected transients. All experiments are performed under ambient conditions.

Figure 2a shows that the magnitude of the light–matter coupling is continuously tunable by means of the control fluence, Φ. The spectra are recorded at a fixed delay, $t_D = 20$ ps, between the near-infrared control and the multi-terahertz probe pulse. In equilibrium ($\Phi = 0$), a single reflectance minimum at $\hbar\omega_c = 113$ meV (top curve, Fig. 2a) attests to the sole resonance of the unexcited cavity, the bare photon mode. With increasing fluence, the system traverses all three regimes of light–matter interaction. Starting with weak coupling ($\Phi \leq 0.03\Phi_0$), Ω_R already exceeds the widths (full-width at half-maximum, ~5 meV) of intersubband and cavity resonances for $\Phi > 0.05\Phi_0$, and two strongly coupled cavity polariton branches are discernible. Further increase of the fluence enhances the separation of the minima to 50 meV, corresponding to a fraction of 44% of the bare photon frequency (Fig. 2b). As discussed in ref. 13, the apparent mode separation is not identical to the vacuum Rabi splitting at the anticrossing point. Only a quantitative simulation of the energy position of the polariton dips (Fig. 2b) allows for extraction of Ω_R. For a correct description of our data, the theory has to go beyond the rotating-wave approximation[14,15]. We include anti-resonant terms in the light–matter Hamiltonian that scale with the ratio $2\Omega_R/\omega_{12}$. These contributions describe the simultaneous creation or annihilation of two excitations with opposite in-plane wavevectors $\mathbf{k}$ and give rise to a two-mode squeezed quantum vacuum[14,15]. By comparison with this theory, we determine that $2\Omega_R = 0.18\omega_{12}$ for our experiment. This value is comparable to the record achieved in delta-doped structures[13] and large enough for the signatures of ultrastrong coupling to be observable[14]. The scheme is expected to be scalable further by means of higher control fluences and a larger number of quantum wells.

The central issue is to explore how rapidly ultrastrong coupling may be activated. Figure 3 displays amplitude spectra recorded at various delay times, t_D ($\Phi = \Phi_0$). For $t_D \leq -50$ fs, the cavity resonance (blue arrow, Fig. 3) shows a minimum amplitude reflectivity below 10%. The control pulse induces dramatic reflectivity changes of order one, on the femtosecond scale. The initial bare photon state is replaced by two coupled polariton branches appearing simultaneously at energy positions of 93 meV and 143 meV (red arrows, Fig. 3). Most notably, the new resonances do not develop by gradual bifurcation of the bare cavity mode as in Fig. 2a. By contrast, switching occurs discontinuously once the control pulse promotes electrons into subband $|1\rangle$.

Immediately following the femtosecond control, the photoexcited charge carriers are in a highly non-equilibrium state which may induce enhanced dephasing of the intersubband transition. A detailed microscopic description of the switching dynamics should thus account for both the quantum kinetic aspects as well as the dynamics of the ultrastrong cavity–intersubband coupling. Notably, for the large coupling strengths achieved in our experiment, we find that dephasing arising from the non-equilibrium nature of the carrier distributions appears to be less important than the dynamics of the cavity polariton splitting. The instantaneous activation of light–matter interaction is

©2009 Macmillan Publishers Limited. All rights reserved

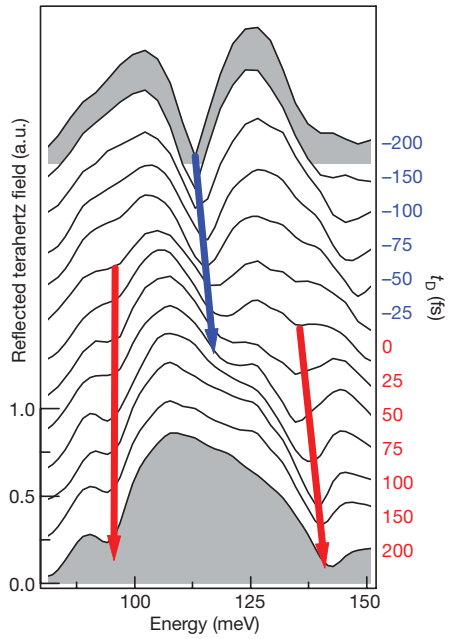


Figure 3 | Non-adiabatic switch-on dynamics of ultrastrongly coupled cavity polaritons. Spectra of the reflected terahertz field are given for various delay times, t_D (vertically off-set). The 12-fs control pulse ($\Phi = 0.1\ \mathrm{mJ\ cm^{-2}}$) arrives at $t_D = 0$. Blue arrow, bare cavity resonance; red arrows, ultrastrongly coupled intersubband cavity polariton branches.

qualitatively well described by the time dependence of the electron distribution in subband $|1\rangle$, as discussed in ref. 14.

Although light–matter coupling is turned on within femtoseconds, Ω_R remains constant on the subsequent nanosecond scale set by the recombination time of the electron–hole pairs. For practical device applications, our scheme may be extended to sub-cycle switch-off: a second infrared control pulse may, for example, promote photo-generated electrons from subband $|1\rangle$ into higher energy levels in the conduction band, hence de-activating the intersubband transitions non-adiabatically. In an alternative, pump–dump, scheme, a pair of identical control pulses (each with pulse area π) may induce strong interband population inversion in subband $|1\rangle$ (switch-on of coupling) followed by ultrafast depopulation through stimulated interband emission (switch-off).

The extreme switching speed demonstrated in Fig. 3 entails unprecedented non-adiabatic phenomena, most strikingly seen in the time domain (Fig. 4). When a few-cycle probe transient (Fig. 4a) impinges on the unexcited modulator, part of its energy is directly reflected from the cavity surface. A second portion is evanescently coupled into the resonator, prepares a coherent photon state, and is subsequently re-emitted (Fig. 4b). This dynamics is encoded in the characteristic twin-pulse structure of the reflected transient (Fig. 4c). The initial burst is due to instantly reflected light whereas the second part results from re-emission. A time-frequency analysis corroborates this scenario. At each point, t, in time, we perform a numerical Fourier transform of the transient in Fig. 4c in a narrow window (width, 100 fs) centred about t. In this way, we map out the instantaneous spectral amplitude $E(t, \omega)$ as a function of time and photon energy (colour plot, Fig. 4c). Filling of the cavity manifests itself in a reflectance minimum at $\hbar\omega = 113$ meV ($t = 50$ fs), whereas subsequent cavity emission causes a delayed spectral peak at the same frequency.

The most intriguing situation arises when we turn light–matter coupling on while a coherent state of bare photons is still present inside the cavity (Fig. 4d): the control pulse (vertical arrow, Fig. 4d) abruptly alters Ω_R during the free cavity decay (second burst of reflected field). Notably, the emission of bare photons is interrupted on a timescale shorter than half an oscillation cycle of light, which is a compelling proof of non-adiabaticity. The subsequent field trace has a characteristic two-mode beating (shown magnified in Fig. 4e; see also Supplementary Fig. 1) and the corresponding spectra (colour plot, Fig. 4d) display minima at energies of 93 meV and 143 meV (indicated by the two diagonal arrows), which are hallmarks of the two polariton branches. Thus, we do not only control the eigenstates of the microcavity, but effectively convert a coherent photon population into ultrastrongly coupled cavity polaritons beating at the

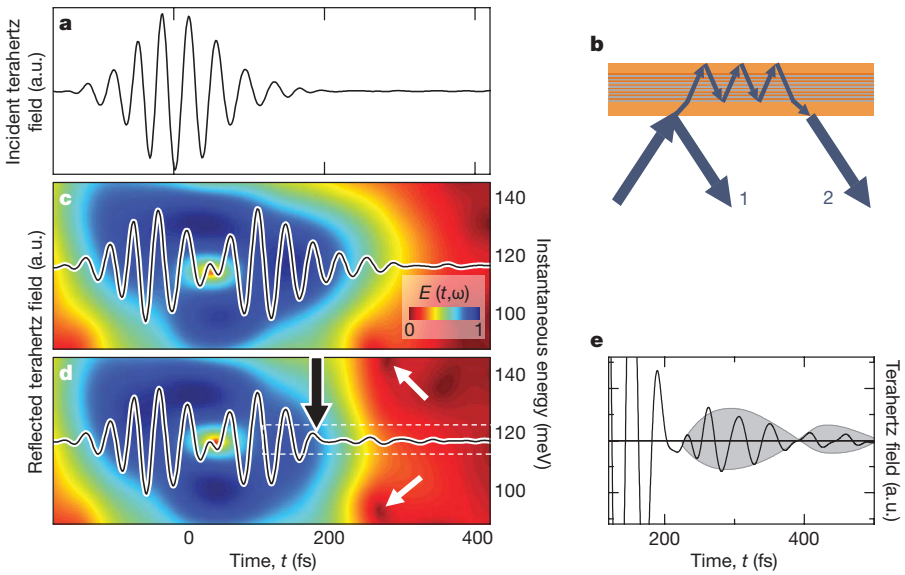


Figure 4 | Perturbed cavity decay. **a**, A measured few-cycle transient incident on the microcavity structure. **b**, Sketch of terahertz reflection from the cavity indicating two distinct contributions: (1) direct reflection at the first mirror of the cavity and (2) re-emission of coherent photon states inside the cavity give rise to a characteristic twin-pulse shape. **c**, Time trace of the experimentally determined terahertz field (black curve) reflected from the unexcited cavity. The corresponding instantaneous amplitude spectrum, $E(t, \omega)$ (colour-coded background: red, low field amplitude; blue, high field amplitude; see colour bar), is obtained as a function of photon energy and time, t, by means of a wavelet transformation of the black curve. **d**, A 12-fs control pulse arrives within the coherence window of the cavity mode (black arrow). The reflected terahertz field (black curve) traces the non-adiabatic switch-on of ultrastrong light–matter coupling: The control pulse abruptly changes the exponential emission decay into a characteristic beating signature, on a sub-cycle scale. The corresponding spectrum (background colour plot) exhibits signatures of both polariton branches, at photon energies of 93 meV and 143 meV (white arrows). **e**, Beating signature (magnified view of dashed rectangle in **d**).

©2009 Macmillan Publishers Limited. All rights reserved

splitting frequency. We suggest the term 'perturbed cavity decay' to describe this non-adiabatic phenomenon.

Our experiments open a new domain of light–matter interaction at the ultrastrong and ultrafast limit. We have observed the dynamics of a coherent photon population during sub-cycle switching. The data provide a benchmark for the most recent, and future, theories of ultrastrong coupling and encourage a systematic search for non-adiabatic QED phenomena, such as the generation of quantum vacuum radiation, reminiscent of the dynamical Casimir effect or Hawking radiation of black holes[20,27,28]. A quantitative estimate using the theory of ref. 15 shows that for the switching times demonstrated (10^{-14} s), the number of vacuum photons released per pulse should be on the order of 10^3, a value that may be measurable with sensitive quantum detectors. Finally, our room-temperature, scalable and tunable semiconductor device, operated with low control fluence, may be highly relevant for next-generation ultrabroadband optical modulators. The concept may be combined with latest developments in robust, sub-10-fs telecommunications-compatible fibre lasers for real-world applications[29].

Received 13 October 2008; accepted 20 January 2009.

1. Raimond, J. M., Brune, M. & Haroche, S. Colloquium: Manipulating quantum entanglement with atoms and photons in a cavity. *Rev. Mod. Phys.* **73,** 565–582 (2001).
2. Brennecke, F. *et al.* Cavity QED with a Bose–Einstein condensate. *Nature* **450,** 268–271 (2007).
3. Weisbuch, C., Nishioka, M., Ishikawa, A. & Arakawa, Y. Observation of the coupled exciton-photon mode splitting in a semiconductor quantum microcavity. *Phys. Rev. Lett.* **69,** 3314–3317 (1992).
4. Dini, D., Köhler, R., Tredicucci, A., Biasiol, G. & Sorba, L. Microcavity polariton splitting of intersubband transitions. *Phys. Rev. Lett.* **90,** 116401 (2003).
5. Reithmaier, J. P. *et al.* Strong coupling in a single quantum dot-semiconductor microcavity system. *Nature* **432,** 197–200 (2004).
6. Chiorescu, I. *et al.* Coherent dynamics of a flux qubit coupled to a harmonic oscillator. *Nature* **431,** 159–162 (2004).
7. Anappara, A. A., Tredicucci, A., Beltram, F., Biasiol, G. & Sorba, L. Tunnel-assisted manipulation of intersubband polaritons in asymmetric coupled quantum wells. *Appl. Phys. Lett.* **89,** 171109 (2006).
8. Dupont, E., Gupta, J. A. & Liu, H. C. Giant vacuum-field Rabi splitting of intersubband transitions in multiple quantum wells. *Phys. Rev. B* **75,** 205325 (2007).
9. Hennessy, K. *et al.* Quantum nature of a strongly coupled single quantum dot-cavity system. *Nature* **445,** 896–899 (2007).
10. Schuster, D. I. *et al.* Resolving photon number states in a superconducting circuit. *Nature* **445,** 515–518 (2007).
11. Sapienza, L. *et al.* Electrically injected cavity polaritons. *Phys. Rev. Lett.* **100,** 136806 (2008).
12. Tsintzos, S. I., Pelekanos, N. T., Konstantinidis, G., Hatzopoulos, Z. & Savvidis, P. G. A. GaAs polariton light-emitting diode operating near room temperature. *Nature* **453,** 372–375 (2008).
13. Anappara, A. A. *et al.* Light-matter excitations in the ultra-strong coupling regime. Preprint at ⟨http://arxiv.org/abs/0808.3720⟩ (2008).
14. Ciuti, C., Bastard, G. & Carusotto, I. Quantum vacuum properties of the intersubband cavity polariton field. *Phys. Rev. B* **72,** 115303 (2005).
15. De Liberato, S., Ciuti, C. & Carusotto, I. Quantum vacuum radiation spectra from a semiconductor microcavity with a time-modulated vacuum Rabi frequency. *Phys. Rev. Lett.* **98,** 103602 (2007).
16. Purcell, E. M. Spontaneous emission probabilities at radio frequencies. *Phys. Rev.* **69,** 681 (1946).
17. Imamoglu, A., Ram, R. J., Pau, S. & Yamamoto, Y. Nonequilibrium condensates and lasers without inversion: Exciton-polariton lasers. *Phys. Rev. A* **53,** 4250–4253 (1996).
18. Kasprzak, J. *et al.* Bose–Einstein condensation of exciton polaritons. *Nature* **443,** 409–414 (2006).
19. Helm, M. in *Intersubband Transitions in Quantum Wells: Physics and Device Applications I* (eds Liu, H. C. & Capasso, F.) 1–99 (Academic, 2000).
20. Yablonovitch, E. Accelerating reference frame for electromagnetic waves in a rapidly growing plasma: Unruh-Davies-Fulling-DeWitt radiation and the nonadiabatic Casimir effect. *Phys. Rev. Lett.* **62,** 1742–1745 (1989).
21. Kübler, C., Huber, R., Tübel, S. & Leitenstorfer, A. Ultrabroadband detection of multi-THz field transients with GaSe electro-optic sensors: approaching the near infrared. *Appl. Phys. Lett.* **85,** 3360–3362 (2004).
22. Ferguson, B. & Zhang, X.-C. Materials for terahertz science and technology. *Nature Mater.* **1,** 26–33 (2002).
23. Tonouchi, M. Cutting-edge terahertz technology. *Nature Photon.* **1,** 97–105 (2007).
24. Huber, R. *et al.* How many-particle interactions develop after ultrafast excitation of an electron–hole plasma. *Nature* **414,** 286–289 (2001).
25. Kröll, J. *et al.* Phase-resolved measurements of stimulated emission in a laser. *Nature* **449,** 698–701 (2007).
26. Gaal, P. *et al.* Internal motions of a quasiparticle governing its ultrafast nonlinear response. *Nature* **450,** 1210–1213 (2007).
27. Unruh, W. G. Second quantisation in the Kerr metric. *Phys. Rev.* **10,** 3194–3205 (1974).
28. Hawking, S. W. Black hole explosions? *Nature* **248,** 30–31 (1974).
29. Sell, A. *et al.* Field-resolved detection of phase-locked infrared transients from a compact Er:fiber system tunable between 55 and 107 THz. *Appl. Phys. Lett.* **93,** 251107 (2008).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements The authors wish to thank C. Kübler, S. Leinß, and D. Seletskiy for assistance at an early stage of this experiment, and I. Carusotto for discussions. Support by the Deutsche Forschungsgemeinschaft through the Emmy Noether Program and SFB767 is gratefully acknowledged.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to R.H. (rupert.huber@uni-konstanz.de).

©2009 **Macmillan Publishers Limited. All rights reserved**

LETTERS

Transparent dense sodium

Yanming Ma[1,2], Mikhail Eremets[3], Artem R. Oganov[2,4]†, Yu Xie[1], Ivan Trojan[3], Sergey Medvedev[3], Andriy O. Lyakhov[2]†, Mario Valle[5] & Vitali Prakapenka[6]

Under pressure, metals exhibit increasingly shorter interatomic distances. Intuitively, this response is expected to be accompanied by an increase in the widths of the valence and conduction bands and hence a more pronounced free-electron-like behaviour. But at the densities that can now be achieved experimentally, compression can be so substantial that core electrons overlap. This effect dramatically alters electronic properties from those typically associated with simple free-electron metals such as lithium (Li; refs 1–3) and sodium (Na; refs 4, 5), leading in turn to structurally complex phases[6–8] and superconductivity with a high critical temperature[9–11]. But the most intriguing prediction—that the seemingly simple metals Li (ref. 1) and Na (ref. 4) will transform under pressure into insulating states, owing to pairing of alkali atoms—has yet to be experimentally confirmed. Here we report experimental observations of a pressure-induced transformation of Na into an optically transparent phase at ~200 GPa (corresponding to ~5.0-fold compression). Experimental and computational data identify the new phase as a wide bandgap dielectric with a six-coordinated, highly distorted double-hexagonal close-packed structure. We attribute the emergence of this dense insulating state not to atom pairing, but to *p–d* hybridizations of valence electrons and their repulsion by core electrons into the lattice interstices. We expect that such insulating states may also form in other elements and compounds when compression is sufficiently strong that atomic cores start to overlap strongly.

Sodium adopts the body-centred cubic (b.c.c.) structure at ambient conditions. Under pressure, it transforms to the face-centred cubic (f.c.c.) structure at 65 GPa (ref. 12) and to the cI16 structure at 103 GPa (refs 7, 13). In recent experiments[7,8,13], further compression to 160 GPa yielded a number of phases in a narrow pressure–temperature range near the minimum of the melting curve (melting temperature $T_{melt} \approx 300$ K at 118 GPa). Theoretical calculations[4,14] suggest stable body-centred tetragonal (Cs-IV) and α-Ga (*Cmca*) structures for pressures above 190 GPa. The α-Ga structure is especially interesting, because the calculations suggest a zero bandgap above 800 GPa owing to pairing of Na atoms[4]. To explore the possible existence of stable yet hitherto unsuspected high-pressure structures and thus advance our understanding of the pressure-induced metal–insulator transition in Na, we undertook an extensive experimental and theoretical study into insulating states in this archetypal metal at high pressures.

Our diamond-anvil cell experiments (see Methods for details) yielded X-ray diffraction data at 101 and 113 GPa corresponding to the known f.c.c. ($a = 3.46$ Å) and cI16 phases ($a = 5.44$ Å), respectively, in agreement with available experimental data[7,12,15]. The corresponding Raman spectra show no features at pressures below 130 GPa. But a pronounced Raman spectrum appears at higher pressures (Fig. 1a), indicating a major phase transformation that according to visual observation is associated with a gradual decrease in the reflection of visible light from the sample. The Raman spectra appearing around 130 GPa are in good accordance with the theoretical spectra calculated for the experimentally observed[7,8] oP8 (*Pnma*) phase (Supplementary Fig. 1d). Above 150 GPa, the Raman spectra again show marked changes, including a strong decrease in intensity, that signify another phase transition (Fig. 1a). The X-ray diffraction pattern of this phase is consistent with the tI19 structure[7,8].

Na becomes optically transparent at pressures of ~200 GPa. (The onset of transparency occurred at 208 GPa in the run shown in Fig. 2a, and at 194 GPa in the run shown in Supplementary Fig. 1a.) The edge in the absorption spectrum (Supplementary Fig. 1c) gives for transparent Na a bandgap of at least ~1.3 eV. The onset of transparency coincides with dramatic changes in the Raman spectra (Fig. 1a), particularly the appearance of a single intense line centred at ~340 cm^{-1}. On releasing the pressure, the transparent phase persists to 182 GPa; at that point, the sample reverts to opaque and exhibits

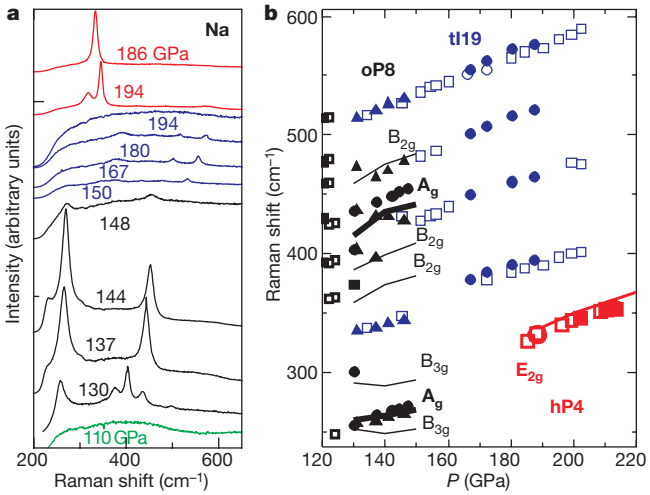


Figure 1 | Raman spectra of sodium. a, Raman spectra obtained at increasing pressures. Spectra of cI16, oP8, tI19 and transparent phases, where the intensities of different phases are representative, are shown by green, black, blue and red lines, respectively. **b**, Pressure dependence of Raman peaks for five pressure runs. The colours correspond to the spectra in **a**. Filled and open symbols correspond to runs on pressure increase and decrease, respectively. Black lines are theoretical Raman modes for the oP8 structure (thicker lines and bold font labels for intense peaks); the red line is the E_{2g} mode for transparent Na-hP4. No Raman calculations were performed for the complex tI19 phase. Phase names are also shown bold.

[1]National Laboratory of Superhard Materials, Jilin University, Changchun 130012, China. [2]Laboratory of Crystallography, Department of Materials, ETH Zurich, Wolfgang-Pauli-Str. 10, CH-8093 Zurich, Switzerland. [3]Max-Planck-Institut für Chemie, Postfach 3060, 55020 Mainz, Germany. [4]Geology Department, Moscow State University, 119992 Moscow, Russia. [5]Data Analysis and Visualization Services, Swiss National Supercomputing Centre (CSCS), Cantonale Galleria 2, 6928 Manno, Switzerland. [6]Consortium for Advanced Radiation Sources, University of Chicago, Chicago, Illinois 60637, USA. †Present address: Department of Geosciences and New York Center for Computational Science, Stony Brook University, Stony Brook, New York 11794-2100, USA.

©2009 Macmillan Publishers Limited. All rights reserved

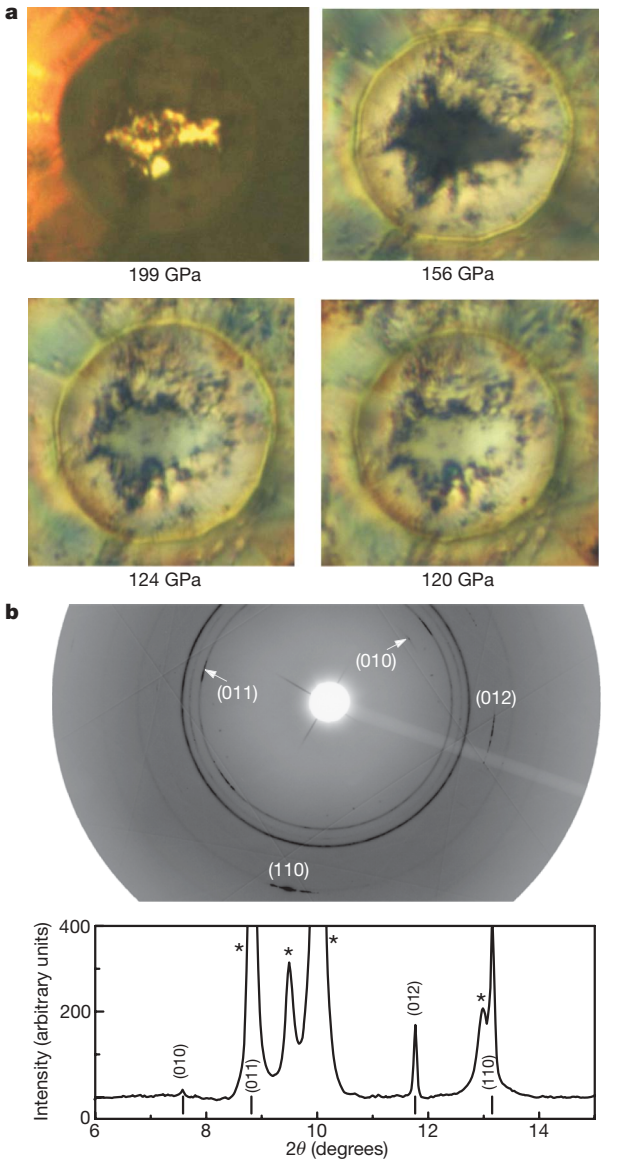


Figure 2 | Phase transformations in Na at megabar pressures. **a**, Photographs of the Na sample taken under combined transmitted and reflected illumination (see also Supplementary Fig. 1) at releasing pressure for different phases: 199 GPa (transparent phase), 156 GPa (tI19), 124 GPa (oP8) and 120 GPa (cI16). The transparent phase was created at 208 GPa, but remained when pressure was dropped to 199 GPa, similar to the outcome of another run (Supplementary Fig. 1). The diameter of the culet of the diamond anvil is 30 μm. **b**, X-ray CCD image and integrated diffraction pattern taken at 190 GPa from the transparent Na sample. Tick marks show the calculated reflection positions for Na-hP4. Asterisks indicate reflections from the Re gasket.

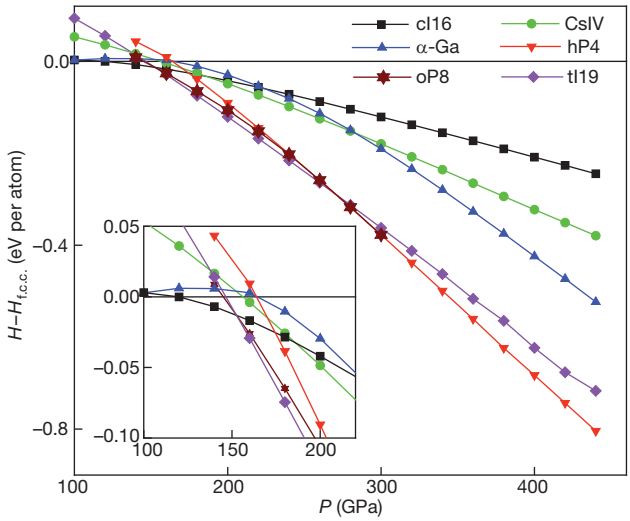


Figure 3 | Enthalpy curves (relative to f.c.c.) as a function of pressure for cI16, CsIV, α-Ga, oP8, tI19 and hP4 structures of Na. The inset shows the transition sequence f.c.c. → cI16 → tI19 (oP8) → hP4 in more detail. Two phase transitions, cI16 → tI19 → hP4, are predicted at 152 and 260 GPa, respectively. The oP8 structure is less stable than the tI19 phase in the pressure range 152–260 GPa, but is energetically very competitive. It is noteworthy that temperature effects (or lattice energies) are not considered in the calculation, but might play an important structural role. Note also that tI19 is an incommensurate structure; thus, its enthalpy calculations were performed with the periodic approximant (Supplementary Fig. 4). The continuous merging of the enthalpy curves of oP8 and hP4 at 250 GPa is due to the second-order nature of the oP8→hP4 transition (Supplementary Fig. 2).

Raman spectra typical of the tI19 phase (Fig. 1a). On further decompression, Na first transforms at 132 GPa to the oP8 phase and finally, at 120 GPa, to a highly reflective metallic state devoid of any Raman signal.

We performed extensive structure searches, which were unbiased by prior knowledge and based on global optimization using the *ab initio* evolutionary methodology for crystal structure prediction[16–19]. Our variable-cell simulations yielded below 260 GPa the experimentally known f.c.c[7,12], cI16[7,8,15] and oP8[7,8] phases (Fig. 3), though the latter is calculated to be marginally less stable than the tI19 phase (see also Fig. 3 inset). Over the pressure range of nearly equal enthalpies for oP8 and tI19, 152–260 GPa, the energy landscape of Na is overall flat yet exhibits a complex pattern of many shallow local minima; as a result, our simulated structure search identified many distinct structures having enthalpies within 30 meV per atom of the ground state. Note, however, that the experiments were carried out at room temperature, whereas the calculations assume $T = 0$ K; moreover, we neglect phonon enthalpies arising from coupling between nuclei and electrons, which can influence the stable structure of the system.

Higher-pressure simulations, at 320 and 1,000 GPa, revealed a simple but unusual structure, which can be described as a double-hexagonal close-packed (d.h.c.p.) structure squeezed along the *c*-axis (Pearson symbol hP4). The 'squeezing' results in an extremely small c/a ratio of, for example, 1.391 at 320 GPa; this ratio is less than half the ideal value of $c/a = 2(8/3)^{1/2} = 3.266$ seen in the normal d.h.c.p structure that corresponds to the densest packing of spherical particles. Atoms in this dense Na structure are six-fold coordinated (Fig. 4a). The characteristic double-well energy profile (Supplementary Fig. 8a) and the six-coordinate topology render Na-hP4 structurally, energetically and electronically distinct from normal d.h.c.p. Remarkably, the energy barrier between two minima in Na-hP4 occurs at $c/a = 3.266$, where Cs-d.h.c.p. has a stable minimum (Supplementary Fig. 8).

The electronic structure of Na-hP4 (Fig. 4b) gives rise to an intriguing insulating state characterized by a large occupancy of *d* orbitals that are involved in *p*–*d* hybridizations. GW calculations (known to provide accurate description of bandgaps, within 10% of experimental values[20]) summarized in Fig. 4d show that Na-hP4 possesses a large energy gap of 1.3 eV at 200 GPa (where the atomic volume is 7.88 Å^3 per atom), and that the gap rapidly increases with increasing pressure to reach 6.5 eV at 600 GPa (5.03 Å^3 per atom).

Although we have been unable to directly determine the structure of the transparent phase observed above pressures of about 200 GPa, the comparison of experimental and theoretical data identifies it reliably as Na-hP4. For example, the optical transparency of the dense Na sample is in agreement with the large bandgaps calculated for this phase. And although the quality of the X-ray diffraction pattern obtained for the transparent phase is insufficient to derive an exact structure solution, the observed reflections are well indexed as Na-hP4, with lattice parameters of $a = b = 2.92$ Å and $c = 4.27$ Å at

©2009 Macmillan Publishers Limited. All rights reserved

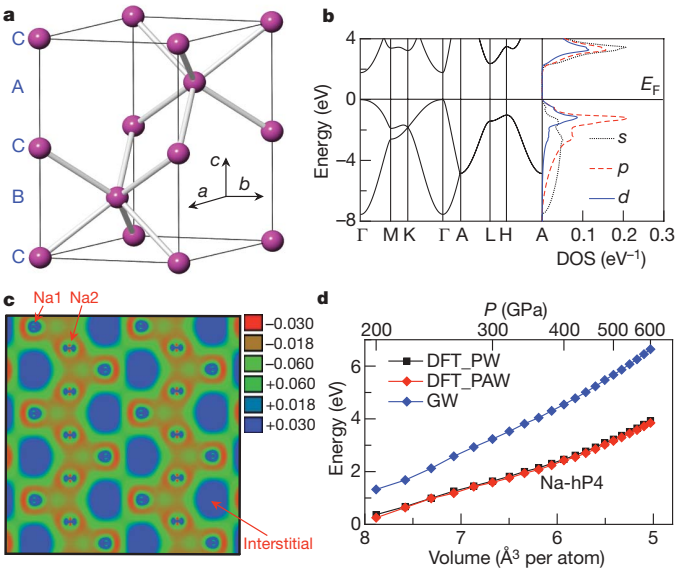


Figure 4 | Structural and electronic properties of Na-hP4. a, Crystal structure of Na-hP4 (space group $P6_3/mmc$). Lattice parameters at 320 GPa (6.50 Å^3 per atom): $a = b = 2.784$ Å and $c = 3.873$ Å. There are two inequivalent atomic positions: Na1 at 2*a* (0.0, 0.0, 0.0) and Na2 at 2*d* (2/3, 1/3, 1/4). **b**, Band structure (left panel) and partial densities of states (DOS; right panel) at 300 GPa (6.67 Å^3 per atom). Though Fermi energy (E_F) has no physical meaning for a nonmetal, we set the highest valence band energy as E_F to indicate that, in analogy with the situation encountered with metals, all bands below E_F are occupied. **c**, Difference charge density (e Å^{-3}; crystal density minus superposition of isolated atomic densities) plotted in the (110) plane at 320 GPa (6.50 Å^3 per atom). **d**, Bandgaps calculated by DFT with the PAW potentials using VASP code (DFT_PAW, red diamonds) and the pseudopotential method using ABINIT code (DFT_PW, black squares), and the GW bandgap (GW, blue diamonds) as a function of volumes for Na-hP4. The upper axis represents theoretical pressure.

190 GPa (Fig. 2b). The position of the observed Raman peak and its dependence on pressure (Fig. 1b) agree well with theoretical calculations for Na-hP4, which possesses only one Raman-active E_{2g} mode. Enthalpy calculations (Fig. 3) confirm the stability of Na-hP4 beyond the tI19 phase. Note that the tI19→hP4 phase transition is predicted to occur at a pressure of 260 GPa (Fig. 3) yet is observed at pressures of ~200 GPa. This discrepancy may in part be due to the athermal nature of our calculations (that is, calculations at 0 K will neglect all entropic contributions), but we expect that the main reason is the well-known overstabilization of metallic states by density functional theory (DFT) calculations (see, for example, ref. 19).

Our calculations reveal that an insulating electronic state emerges because compression causes the 3*d* bands to rapidly drop in energy relative to the 3*p* bands and increasingly hybridize with them[4]. This hybridization is the key to strong electron localization: a marked charge accumulation occurs only in the open interstitial regions (Fig. 4c, Supplementary Figs 6 and 7). The shortest Na–Na distance in Na-hP4 is 1.89 Å at 300 GPa (6.67 Å^3 per atom), which implies strong core–valence overlap (the 3*s* and 2*p* orbital radii in Na are 1.71 and 0.28 Å, respectively) and even significant core–core overlap between neighbouring Na atoms (the ionic radius of Na^+ is 1.02 Å). A consideration of the geometry of charge maxima relative to the atoms implies *pd* and pd^2 hybridizations for Na1 in the octahedral sites and Na2 in the triangular prismatic coordination, respectively—consistent with the site-*l*-projected density of states (Fig. 4b and Supplementary Fig. 10). Thus, the insulating state in Na-hP4 arises from the strong localization of valence electrons in the interstices of Na-hP4 (Fig. 4c and Supplementary Fig. 6) enabled by *p*–*d* hybridization. In this context, it is interesting to note that in contrast to the *p*–*d* hybridization in Na, calculations[3] for other alkalis suggest they are dominated either by *p* electrons (Li) or by *d* electrons (K, Rb and Cs). Na-hP4 might thus be viewed as the intermediate structure between the six-coordinated structure predicted for Li and the ideal d.h.c.p. structures of K, Rb and Cs (ref. 3).

Note that while the stacking of close-packed layers of Na atoms is CACBCACB… (underlining indicates layers that contain interstitial electron density maxima; Fig. 4a) as in any d.h.c.p. structure, the interstitial charge density maxima are located only in layers A and B (that is, in every other layer) and form a nearly perfect h.c.p arrangement (layer stacking of ABAB… and $c/a \approx 1.3$–1.6, close to the ideal $c/a = 1.633$ for the h.c.p. structure). This finding explains the approximate halving of the c/a ratio as a reflection of the close packing of the interstitial electron density maxima, rather than Na atoms. This suggests that repulsion between the interstitial electron pairs is a major structure-forming interaction in Na-hP4.

As the number of ionic cores (each stripped of one valence electron) is exactly twice that of interstitial electron density maxima, Na-hP4 is an analogue of Ni_2In-type structures (considered also in ref. 7): the ionic cores form the Ni sublattice and the interstitial density maxima the In sublattice. And as pressure-induced transitions in the A_2X compounds[21] lead to an increase of the coordination number of X sites (that is, voids in Na-hP4), pressure-induced transitions in Na lead to an increase of the coordination number of lattice interstices: the coordination numbers are 8, 9 and 11, respectively, in the cI16, oP8 and hP4 phases of Na.

The case of dense Na provides an unexpectedly extreme example of a wide-gap insulator created by compression of a 'simple' metal. This novel dielectric consists of ionic cores and localized interstitial electron pairs, in analogy to electrides[22] where the interstitial electron density maxima play the role of anions. Dense insulating states such as reported here for Na may also be favoured in other elements and compounds when atomic cores are made to strongly overlap.

METHODS SUMMARY

For the experiments, sodium was loaded into a diamond-anvil cell in a nitrogen atmosphere and Raman spectra obtained with a 25 mW HeNe laser. X-ray studies were conducted at the Advanced Photon Source at the Argonne National Laboratory, USA. The evolutionary structure search used DFT for *ab initio* structure calculations.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 18 June 2008; accepted 19 January 2009.

1. Neaton, J. B. & Ashcroft, N. W. Pairing in dense lithium. *Nature* **400,** 141–144 (1999).
2. Tamblyn, I., Raty, J. Y. & Bonev, S. A. Tetrahedral clustering in molten lithium under pressure. *Phys. Rev. Lett.* **101,** 075703 (2008).
3. Ma, Y., Oganov, A. R. & Xie, Y. High pressure structures of lithium, potassium, rubidium predicted by ab initio evolutionary algorithm. *Phys. Rev. B* **78,** 014102 (2008).
4. Neaton, J. B. & Ashcroft, N. W. On the constitution of sodium at higher densities. *Phys. Rev. Lett.* **86,** 2830–2833 (2001).
5. Raty, J. Y., Schwegler, E. & Bonev, S. A. Electronic and structural transitions in dense liquid sodium. *Nature* **449,** 448–451 (2007).
6. Hanfland, M., Syassen, K., Christensen, N. E. & Novikov, D. L. New high-pressure phases of lithium. *Nature* **408,** 174–178 (2000).
7. Hanfland, M., Syassen, K., Loa, L., Christensen, N. E. & Novikov, D. L. Na at megabar pressures. *Poster at 2002 High Pressure Gordon Conference* (2002).
8. Gregoryanz, E. *et al.* Structural diversity of sodium. *Science* **320,** 1054–1057 (2008).
9. Shimizu, K., Ishikawa, H., Takao, D., Yagi, T. & Amaya, K. Superconductivity in compressed lithium at 20 K. *Nature* **419,** 597–599 (2002).
10. Struzhkin, V. V., Eremets, M. I., Gan, W., Mao, H. K. & Hemley, R. J. Superconductivity in dense lithium. *Science* **298,** 1213–1215 (2002).
11. Deemyad, S. & Schilling, J. S. Superconducting phase diagram of Li metal in nearly hydrostatic pressures up to 67 GPa. *Phys. Rev. Lett.* **91,** 167001 (2003).
12. Hanfland, M., Loa, I. & Syassen, K. Sodium under pressure: bcc to fcc structural transition and pressure-volume relation to 100 GPa. *Phys. Rev. B* **65,** 184109 (2002).
13. Gregoryanz, E., Degtyareva, O., Somayazulu, M., Hemley, R. J. & Mao, H. K. Melting of dense sodium. *Phys. Rev. Lett.* **94,** 185502 (2005).
14. Christensen, N. E. & Novikov, D. L. High-pressure phases of the light alkali metals. *Solid State Commun.* **119,** 477–490 (2001).
15. McMahon, M. I. *et al.* Structure of sodium above 100 GPa by single-crystal x-ray diffraction. *Proc. Natl Acad. Sci. USA* **104,** 17297–17299 (2007).

©2009 Macmillan Publishers Limited. All rights reserved

16. Glass, C. W., Oganov, A. R. & Hansen, N. USPEX—Evolutionary crystal structure prediction. *Comput. Phys. Commun.* **175,** 713–720 (2006).
17. Oganov, A. R. & Glass, C. W. Crystal structure prediction using ab initio evolutionary techniques: Principles and applications. *J. Chem. Phys.* **124,** 244704 (2006).
18. Oganov, A. R., Glass, C. W. & Ono, S. High-pressure phases of $CaCO_3$: Crystal structure prediction and experiment. *Earth Planet. Sci. Lett.* **241,** 95–103 (2006).
19. Ma, Y., Oganov, A. R. & Glass, C. W. Structure of the metallic ζ-phase of oxygen and isosymmetric nature of the ε-ζ phase transition: Ab initio simulations. *Phys. Rev. B* **76,** 064101 (2007).
20. Shishkin, M. & Kresse, G. Self-consistent GW calculations for semiconductors and insulators. *Phys. Rev. B* **75,** 235102 (2007).
21. Leger, J. M. & Haines, J. Crystal chemistry of the AX_2 compounds under pressure. *Eur. J. Solid State Inorg. Chem.* **34,** 785–796 (1997).
22. Dye, J. L. Electrides: From 1D Heisenberg chains to 2D pseudo-metals. *Inorg. Chem.* **36,** 3816–3826 (1997).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank the Swiss National Science Foundation (grants 200021-111847/1 and 200021-116219), CSCS and ETH Zurich for the use of supercomputers. Parts of the calculations were performed on the Skif supercomputer (Moscow State University, Russia) and at the Joint Supercomputer Centre of the Russian Academy of Sciences (Moscow). We acknowledge partial support from DFG (grants Er 539/1/2-1) and the China 973 Program (no. 2005CB724400). Part of the experimental work was performed at GeoSoilEnviroCARS (Sector 13), Advanced Photon Source (APS), Argonne National Laboratory.

Author Contributions Y.M. proposed the research and predicted the new structures. Y.M., Y.X. and A.R.O. did the calculations. M.E., I.T., S.M. and V.P. performed the experiments. Y.M., A.R.O. and M.E. contributed substantially to data interpretation and wrote the paper. A.L. wrote the latest version of the structure prediction code, and M.V. helped in data analysis. Y.M, M.E. and A.R.O contributed equally to this paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to Y.M. (mym@jlu.edu.cn).

©2009 **Macmillan Publishers Limited. All rights reserved**

METHODS

High-pressure experiments. Our diamond-anvil cell (DAC) experiments used diamonds with 30–50 μm culet bevelled to 8–9° and gaskets of rhenium or tungsten, or of cubic BN (cBN) powder mixed with epoxy. Sodium of 99.95% purity (Alfa Aesar) was loaded in a glove box in an atmosphere of pure nitrogen containing <0.1 p.p.m. oxygen and water. Before loading, the DAC was kept at 130 °C in a vacuum for 20 h to remove traces of water and gases at the surface of the gasket, thereby excluding reaction of the clamped sodium with rhenium or cBN/epoxy gaskets. Pressure was measured mainly by the Raman shift of the high-frequency edge of the stressed diamonds[23], and was in good agreement with the pressure obtained from X-ray diffraction data for Re and cBN in checking experiments. We estimate the pressure error to be within 5 GPa at the highest pressures, and less at lower pressures. An extremely sharp Raman edge of the diamond anvil indicated an excellent hydrostaticity of Na up to the highest pressures. An HeNe laser of power 25 mW was used for excitation of Raman spectra. Angle dispersive X-ray diffraction studies were performed at station 13-IDD at the Advanced Photon Source, Argonne National Laboratory, USA ($\lambda = 0.3344$ Å).

Evolutionary structure search and energy calculations. The evolutionary structure search was performed using the USPEX code[16–18], and the underlying *ab initio* structure relaxations were performed using DFT within the generalized gradient approximation (GGA)[24] and the frozen-core all-electron projector-augmented wave (PAW)[25,26] method, as implemented in the VASP code[27]. The quasiparticle energies were calculated within the GW approximation, as implemented in the ABINIT code[28], in which self-energy corrections are added to the Kohn–Sham energies at the selected k points that characterize the DFT bandgap.

We treated the 2*s*, 2*p* and 3*s* electrons as valence in the PAW potential adapted from the VASP library and used the plane-wave kinetic energy cut-off of 910 eV, which gave excellent convergence of the total energies, energy differences and structural parameters. We used the Monkhorst–Pack k meshes of $16 \times 16 \times 16$, $12 \times 12 \times 12$, $19 \times 19 \times 13$, $16 \times 16 \times 12$, $12 \times 16 \times 12$, $6 \times 6 \times 6$ and $20 \times 20 \times 14$ for f.c.c., cI16, α-Ga, Cs-IV, oP8, tI19 and hP4, respectively, which gave excellent convergence of the total energies (within 1 meV per atom) in the enthalpy calculations. GW calculations for Na-hP4 were performed with 131 k points in the first Brillouin zone, and an energy cut-off of 20 Ry was chosen for calculation of the Coulomb matrix. In the calculation of the self-energy matrix, 18 occupied bands and 120 unoccupied bands were explicitly treated. With the choice of these parameters, the bandgaps were found to converge within 0.01 eV.

23. Eremets, M. I. Megabar high-pressure cells for Raman measurements. *J. Raman Spectrosc.* **34**, 515–518 (2003).
24. Perdew, J. P., Burke, K. & Ernzerhof, M. Generalized gradient approximation made simple. *Phys. Rev. Lett.* **77**, 3865–3868 (1996).
25. Blöchl, P. E. Projector augmented-wave method. *Phys. Rev. B* **50**, 17953–17979 (1994).
26. Kresse, G. & Joubert, D. From ultrasoft pseudopotentials to the projector augmented-wave method. *Phys. Rev. B* **59**, 1758–1775 (1999).
27. Kresse, G. & Furthmüller, J. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Phys. Rev. B* **54**, 11169–11186 (1996).
28. Gonze, X. *et al*. First-principles computation of material properties: The ABINIT software project. *Comput. Mater. Sci.* 25, 478–492 (2002); ⟨http://www.abinit.org/⟩.

©2009 Macmillan Publishers Limited. All rights reserved

LETTERS

Direct observation of a pressure-induced metal-to-semiconductor transition in lithium

Takahiro Matsuoka[1] & Katsuya Shimizu[1]

Lithium, the lightest metal, has long been considered to have a 'simple' electronic structure that can be well explained within the nearly-free-electron model. But lithium does not stay 'simple' under compression: rather than becoming more free-electron-like as pressure is increased, first-principles calculations[1,2] suggest that it transforms into a semi-metal or semiconductor. Experimentally, it has been shown that dense lithium adopts low-symmetry structures[3,4]; there is also evidence that its resistivity increases with pressure[5–8]. However, the electronic transport properties of lithium have so far not been directly monitored as a function of increasing static pressure. Here we report electrical resistance measurements on lithium in a diamond anvil cell up to pressures of 105 GPa, which reveal a significant increase in electrical resistivity and a change in its temperature dependence near 80 GPa. Our data thus provide unambiguous experimental evidence for a pressure-induced metal-to-semiconductor transition in a 'simple' metallic element.

Ever since it was predicted[1,2] that the light alkali metal lithium (Li) may form pairs of atoms ('dimers') and transform to a semi-metal or semiconductor in its highly compressed state near 100 GPa, much effort has been directed to the challenge of studying dense Li. An X-ray diffraction study showed[3] that face-centred-cubic (f.c.c.) Li becomes unstable and transforms to a cubic crystalline phase (*I*-43*d*) with a 16-atom unit cell at around 40 GPa, confirming the strong preference of dense lithium for low-symmetry structures. A more recent experimental investigation[4] identified a possible structural transition near 80 GPa, but was unable to clarify the structure of the higher-pressure phase. Electrical resistance measurements have supported the view that dense Li does not behave as a 'simple' metal, by documenting a continuous and significant resistivity increase with pressure under static compression to about 30 GPa (refs 5, 6) and under shock-compression to about 60 GPa (ref. 7) and 180 GPa (ref. 8). The visual observation that some dense Li phases are dark[9,10] also indicates loss of metallicity. A theoretical calculation predicted a three-coordinated orthorhombic *Cmca*-24 structure[11] to be stable from 88 GPa to at least 140 GPa, in agreement with a recent structural search[12] that concluded that *Cmca*-24 is stable up to 300 GPa and then transforms into a cubic $P4_132$ structure[12] containing six-fold coordinated lithium. However, these structures were all predicted to be metallic, and their structure cannot be reconciled with the observation[10] of a Raman vibrational band in Li above 70 GPa to 120 GPa that indicates the existence of a broken symmetry phase.

We performed electrical resistance measurements on Li in a diamond anvil cell (DAC) at pressures up to 105 GPa, which exceeds the pressure reached in all previous studies employing static compression. Images and schematic illustrations of the experimental set-up are shown in Fig. 1. For all pressures above 8 GPa, experiments were performed in a refrigerator below 50 K to prevent pressure-induced diffusion of Li into the diamond anvils[3,4,10,13], which has been one of the main problems preventing investigations of the physical properties of Li under dense conditions.

The present electrical resistance experiments at pressures as high as 105 GPa were made possible by using an improved insulating gasket material and advanced fabrication methods to deposit fine electrodes on the diamond anvil surface facing the sample. Regarding the first technical advance, we note that electrical resistance measurements are only possible when the gasket surface touching the sample and

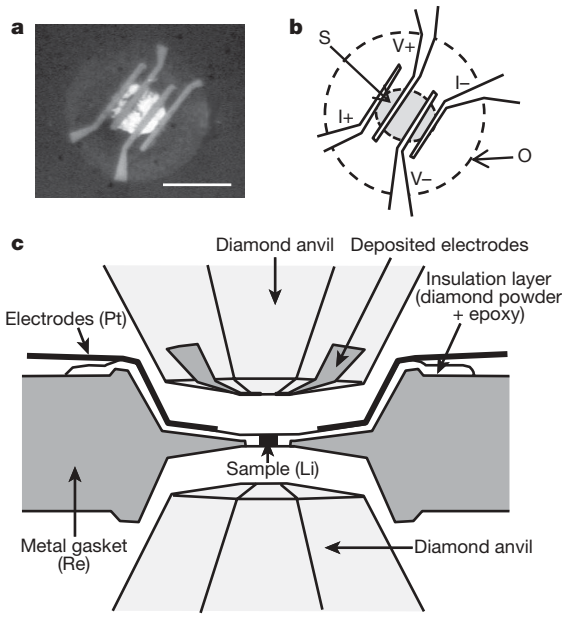


Figure 1 | Experimental set-up. a, Image of the sample and electrodes placed on the insulating gasket, looking through the diamond anvil. Scale bar, 50 μm. The electrodes consist of a bottom layer of titanium (Ti) and a top layer of copper (Cu) and are prepared such that they contact the sample at the centre of the gasket. **b**, Schematic drawing of the components seen in **a**, illustrating how four-point-contact measurements were performed. The deposited electrodes provide a contact between the sample (S) and platinum thin-foil electrodes, which are placed on the insulating layer and extended outside the sample chamber. The electrical resistance was measured by a resistance bridge. The dashed circle (O) indicates the outline of the 100 μm culet. The electrode pair I+ and I− carries current, and the pair V+ and V− senses voltage. **c**, Schematic cross-section of the experimental set-up in the diamond anvil cell (DAC). Li (^{7}Li, 99.95% purity, Johnson-Matthey) was loaded under an Ar atmosphere with less than 1 p.p.m. oxygen and moisture. No pressure medium was used. Li and the electrodes are insulated from the metal gasket by an insulation layer made of a compressed mixture of diamond powder and epoxy. Pressure medium was not used. The shift of the first-order Raman band spectra of the diamond anvil facing the sample was used, with a proposed calibration[21], to estimate pressure. Pressure was controlled at 25 K by a diamond anvil cell driven by He gas. Samples were compressed at room temperature to some pressure below 4 GPa, and then cooled to the experimental temperature.

[1]KYOKUGEN, Center for Quantum Science and Technology under Extreme Conditions, Osaka University, 1-3 Machikaneyama, Toyonaka, Osaka 560-8531, Japan.

©2009 Macmillan Publishers Limited. All rights reserved

electrodes is an insulator. We usually use insulating gaskets made of Al_2O_3 or cubic BN, but the very soft metal Li can easily escape from the sample chamber under these conditions. Previous studies up to 50 GPa therefore used as sample chamber a pit machined into the top diamond surface[13]. We found that when a compressed mixture of diamond powder and epoxy is placed as an insulating layer on a metal gasket (Fig. 1c), Li is contained even in the megabar region without deformation of the sample chamber. We found no evidence of contamination due to possible chemical reactions between Li and epoxy. For instance, the electrical resistance and the sample appearance did not change when holding the system for some time at constant pressures and temperatures. The second crucial technical advance involved deposition of sufficiently fine electrodes on the diamond anvil surface facing the sample (Fig. 1a) to enable electrical resistance measurement by the four-point contact method, even on micrometre-sized samples. This was achieved by using photo-lithography and a vapour deposition system to deposit layers of Ti (bottom) and Cu (top) such that the distance between each electrode in the central area (where they contact the sample) is 10 μm, with a layer thickness of less than 1 μm.

In Fig. 2a, resistivity at 25 K is plotted as a function of pressure. (Resistivity was calculated from measured resistance values by using the area/length ratio of the Li sample between the electrodes measured before loading; see Fig. 2a legend for details.) The resistivity at 10 GPa was of the order of 10^{-3} mΩ cm, which indicates a good conductor. It then increased significantly with pressure to reach at 70 GPa a value of 0.2 mΩ cm, which is comparable to the resistivities of the semi-metals Sb (0.04 mΩ cm)[14] and graphite (~1.0 mΩ cm)[14]. This increase is fairly continuous, except for a small drop around 40 GPa. Above 70 GPa, the resistivity showed a dramatic increase by four orders of magnitude. As can be seen from the vertical lines in Fig. 2a that indicate the structural phase boundaries reported[4] in a recent X-ray diffraction study at 25 K (ref. 4), the huge electrical resistivity increase occurs as the *cI*16 phase transforms into higher-pressure phases that we name here Li-VI and Li-VII. (Note that these two high-pressure phases were referred to as hp-I and hp-II in ref. 4; to make for a more consistent naming, we now call these two phases Li-VI and L-VII, respectively, to indicate that they are the fifth and sixth phases of Li observed at high pressure and low temperature, following the known body-centred cubic (b.c.c.), R9 (refs 15–17), f.c.c. (ref. 18), *hR*1 (ref. 3) and *cI*16 (ref. 3) structures known to exist below room temperature. The structures of Li-VI and Li-VII are not yet established.) The small resistivity drop around 40 GPa can be attributed to phase transitions from f.c.c. to *hR*1, and from *hR*1 to *cI*16. The continuity of the resistivity increase between about 70 GPa and 80 GPa, when the system transits to a semiconductor, is interpreted as the result of a phase mixture. Several measurements were carried out and fully reproducible results were obtained in each measurement. And in other measurements, the resistivity curve obtained through pressure unloading traced the curve obtained through pressure loading with little evidence of hysteresis (Supplementary Information).

Figure 2b shows electrical resistivity as a function of temperature at seven different pressures spanning the range shown in Fig. 2a. Below 68 GPa, the resistivity–temperature slope $d\rho/dT$ clearly shows metallic character: a positive temperature dependence. The steep drops in electrical resistivity apparent in the curves obtained for constant pressures of 34, 42 and 68 GPa indicate superconducting transitions, in agreement with previous work[4,13,19,20]. For pressures of 78 GPa and above, there is a significant increase in electrical resistivity and the curves have a negative slope, $d\rho/dT$, characteristic of a semiconductor. The semiconducting character is maintained up to 95 GPa.

The three highest-pressure resistance curves in Fig. 2b indicate a very small activation energy for electrical conduction. The activation energy (E_g) can be estimated using $\rho = \rho_0 \exp(E_g/kT)$, where k is the Boltzmann constant and T is temperature. We found that plots of $\ln(\rho)$ against kT

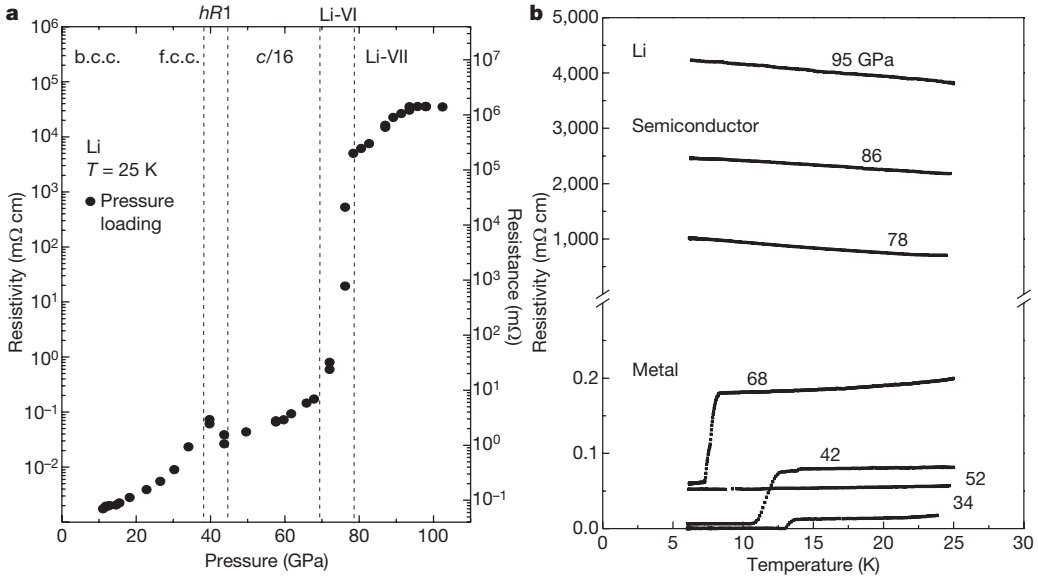


Figure 2 | Electrical resistivity changes with pressure and temperature. **a**, Measured resistance values (right vertical axis) and derived electrical resistivity values (left vertical axis) at 25 K as a function of applied pressure. Vertical broken lines indicate structural phase boundaries at 25 K determined in an earlier synchrotron X-ray diffraction study[4]. All experimental data are obtained during pressure loading cycles. Resistivity is calculated using the area/length ratio measured before loading the sample, for which we use $l \times w \times t$ (10 μm × 45 μm × 40 μm). Here l is the separation between two electrodes for voltage measurement, w is the width of sample perpendicular to current flow, and t is the thickness of sample. According to a relevant equation of state[3], the atomic volume at 50 GPa is about 40% of the atomic volume at ambient conditions. Because visible observation indicated that l and w do not change, we assume that only t changes by compression by 60% at 50 GPa. Systematic uncertainties in electrical resistivity values will then be less than 60% up to pressures of 50 GPa, and more than that above 50 GPa. Although large, the uncertainties allow observation of changes in electrical character. Pressure determination based on diamond Raman vibrations deviates somewhat from the widely used ruby pressure scale[22], by about 10 GPa at 70 GPa. **b**, Temperature dependence of electrical resistivity, between about 5.5 and 25 K. Shown are curves obtained at seven different applied pressures, showing that the slope of the temperature dependence of the electrical resistivity changed from positive to negative at and above approximately 78 GPa. For resistivity curves obtained at lower pressures, steep drops arise from the superconducting transitions of Li (refs 6, 13, 19, 20). For the curve at 52 GPa, the onset of the superconducting transition was not observed within the experimental temperature range.

©2009 Macmillan Publishers Limited. All rights reserved

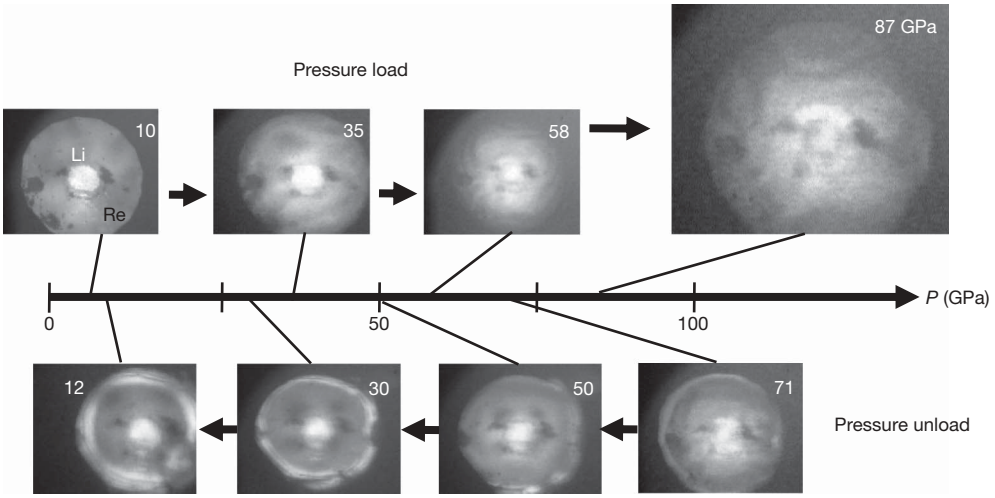


Figure 3 | Visual observation of pressure-induced changes in metallic reflection. Shown are photographs of the lithium sample under pressure, taken by looking through a diamond anvil; number at top right of each image indicates pressure in GPa. Top and bottom rows document changes during pressure loading and unloading, respectively. Note that the images become somewhat obscured by the deformation of the diamond anvil. Li is contained in a sample chamber prepared in the centre of the rhenium metal gasket and retains its shiny metallic reflection at relatively lower pressures (below about 60 GPa). At 87 GPa, most of the central area of the sample has lost a considerable amount of its reflection. Upon pressure release from 87 GPa, the shiny metallic reflection is recovered.

are not perfectly linear and therefore used the tangent of their slopes at 20 K to estimate the activation energy E_g, obtaining values of 0.83 meV ($\rho_0 = 4.72$ (Ω cm)) at 78 GPa, 0.23 meV ($\rho_0 = 19.74$ (Ω cm)) at 86 GPa, and 0.20 meV ($\rho_0 = 35.05$ (Ω cm)) at 95 GPa. These small estimated activation energy values may indicate that high-pressure Li is a zero-gap semiconductor, or may arise due to the effects of impurities, defects, pressure gradients, and so on. Regarding the possible role of impurities (due to chemical reactions of Li with surrounding materials); this was considered but no evidence of contamination was found (see Supplementary Information). However, we note that our experimental set-up limits measurements to a narrow temperature range (spanning 5 K to ~25 K), and that the small activation energy values may thus simply reflect our experimental limitations. More definite values and an indication of the factor(s) that influence activation energies could be obtained by extending the experimental temperature range to higher temperatures such as 200 K, though this will require technical advances to overcome practical difficulties arising from Li diffusion into diamond anvils.

In our measurements, the diamond anvils failed above 105 GPa; this made it impossible to reduce the pressure after observing the metal-to-semiconductor transition and directly probe the reversibility of the transition. But Fig. 3 shows photographs of Li under the same pressure and temperature conditions as used in the electrical resistivity experiments, and document the changes on loading and unloading of the DAC. At pressures exceeding 80 GPa, the metallic reflection from the Li showed a strong reduction while the reflection from the rhenium metal gasket was almost independent of pressure. As pressure was unloaded from 86 GPa to ambient pressure, Li recovered its metallic reflection. These observations of a reversible pressure-induced change from a shiny metal to a dark material are in good agreement with the results of the electrical resistivity experiments.

Taken together, our results clearly show that Li transforms from a metal to a semiconductor near 80 GPa. This provides a definite answer to the question of whether the application of pressure can drive a near-free-electron metal into a dense insulating state. The crystal structure of the semiconductor Li phase is unfortunately not yet known. But a vibrational band has been observed[10] in Li above 70 GPa (that is, close to the pressure where Li transforms from a metal to a semiconductor), and it suggests that the semiconductor phase found in this study has broken symmetry. Clearly, technical advances are needed to enable further experiments to clarify the unusual structural, dynamical and electronic properties of Li and other alkali metals. Extending experimental studies to higher pressures would be particularly attractive, as Li may revert back to a metal. Such studies might possibly also uncover unexpected phenomena such as the emergence of high-temperature superconductivity in even denser states, and more generally advance our understanding of the physical properties of 'simple' metallic elements under high-pressure conditions.

Received 14 May 2008; accepted 19 January 2009.

1. Neaton, J. B. & Ashcroft, N. W. Pairing in dense lithium. *Nature* **400,** 141–144 (1999).
2. Christensen, N. E. & Novikov, D. L. High-pressure phases of the alkali metals. *Solid State Commun.* **119,** 477–490 (2001).
3. Hanfland, M., Syassen, K., Christensen, N. E. & Novikov, D. L. New high-pressure phase of lithium. *Nature* **408,** 174–178 (2000).
4. Matsuoka, T. *et al.* Superconductivity and crystal structure of lithium under high pressure. *J. Phys. Conf. Ser.* **121,** 052003 (2008).
5. Stager, R. A. & Drickamer, H. G. Effect of temperature and pressure on the resistance of four alkali metals. *Phys. Rev.* **132,** 124–127 (1963).
6. Lin, T. H. & Dunn, K. J. High-pressure and low-temperature study of electrical resistance of lithium. *Phys. Rev. B* **33,** 807–811 (1986).
7. Fortov, V. E. *et al.* Anomalous electric conductivity of lithium under quasiisentropic compression to 60 GPa (0.6 Mbar). Transition into a molecular phase? *JETP Lett.* **70,** 628–632 (1999).
8. Bastea, M. & Bastea, S. Electrical conductivity of lithium at megabar pressures. *Phys. Rev. B* **65,** 193104 (2002).
9. Mori, Y., Zha, C. & Ruoff, A. L. in *Science and Technology of High Pressure* (eds Manghani, M. H., Nellis, W. J. & Nicol, M. F.) 421–424 (Univ. Press India, 2000).
10. Goncharov, A. F., Struzhkin, V. V., Mao, H. K. & Hemley, R. J. Spectroscopic evidence for broken-symmetry transitions in dense lithium up to megabar pressures. *Phys. Rev. B* **74,** 184114 (2005).
11. Rousseau, R., Uehara, K., Klug, D. D. & Tse, J. S. Symmetry transition of elemental lithium up to 140 GPa. *ChemPhysChem* **6,** 1703–1706 (2005).
12. Ma, Y., Oganov, A. R. & Xie, Y. High-pressure structures of lithium, potassium, and rubidium predicted by an *ab initio* evolutionary algorithm. *Phys. Rev. B* **78,** 014102 (2008).
13. Shimizu, K., Ishikawa, H., Takao, D., Yagi, T. & Amaya, K. Superconductivity in compressed lithium at 20 K. *Nature* **419,** 597–599 (2002).
14. Cottrell, A. H. *An Introduction to Metallurgy* 2nd ed. (Maney Publishing, London, 1997).
15. Overhauser, A. W. Crystal structure of lithium at 4.2 K. *Phys. Rev. Lett.* **53,** 64–65 (1984).
16. Vaks, V. G. *et al.* An experimental and theoretical study of martensitic phase transitions in Li and Na under pressure. *J. Phys. Condens. Matter* **1,** 5319–5335 (1989).
17. Smith, H. G., Berliner, R., Jorgensen, J. D., Nielsen, M. & Trivisonno, J. Pressure effects on the martensitic transformation in metallic lithium. *Phys. Rev. B* **41,** 1231–1234 (1990).
18. Hanfland, M., Loa, I., Syassen, K., Schwarz, U. & Takemura, K. Equation of state of lithium to 21 GPa. *Solid State Commun.* **112,** 123–127 (1999).
19. Struzhkin, V. V., Eremets, M. I., Gan, W., Mao, H. K. & Hemley, R. J. Superconductivity in dense lithium. *Science* **298,** 1213–1215 (2002).
20. Deemyad, S. & Schilling, J. S. Superconducting phase diagram of Li metal in nearly hydrostatic pressures up to 67 GPa. *Phys. Rev. Lett.* **91,** 167001 (2003).

©2009 Macmillan Publishers Limited. All rights reserved

21. Akahama, Y. & Kawamura, H. High-pressure Raman spectroscopy of diamond anvils to 250 GPa: Method for pressure determination in the multimegabar pressure range. *J. Appl. Phys.* **96,** 3748–3751 (2004).
22. Mao, H. K., Xu, J. & Bell, P. M. Calibration of the ruby pressure gauge to 800 kbar under hydrostatic conditions. *J. Geophys. Res.* **91,** 4673–4676 (1986).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank N. W. Ashcroft and J. S. Schilling for discussions. This work was supported in part by a Grant-in-Aid for Scientific Research (S), 19104009 and Global COE Program (Core Research and Engineering of Advanced Materials-Interdisciplinary Education Center for Materials Science), MEXT, Japan, and a Grant-in-Aid for JSPS Fellows (19•52753). We acknowledge J. Tse, Y. Yao and D. Klug for discussions and suggestions.

Author Contributions T. M. and K. S. performed the experiments and analysed the data. Both authors wrote the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to T.M. (matsuoka@djebel.mp.es.osaka-u.ac.jp).

©2009 **Macmillan Publishers Limited. All rights reserved**

Vol 458|12 March 2009|doi:10.1038/nature07853

LETTERS

Battery materials for ultrafast charging and discharging

Byoungwoo Kang[1] & Gerbrand Ceder[1]

The storage of electrical energy at high charge and discharge rate is an important technology in today's society, and can enable hybrid and plug-in hybrid electric vehicles and provide back-up for wind and solar energy. It is typically believed that in electrochemical systems very high power rates can only be achieved with supercapacitors, which trade high power for low energy density as they only store energy by surface adsorption reactions of charged species on an electrode material[1–3]. Here we show that batteries[4,5] which obtain high energy density by storing charge in the bulk of a material can also achieve ultrahigh discharge rates, comparable to those of supercapacitors. We realize this in $LiFePO_4$ (ref. 6), a material with high lithium bulk mobility[7,8], by creating a fast ion-conducting surface phase through controlled off-stoichiometry. A rate capability equivalent to full battery discharge in 10–20 s can be achieved.

Like any lithium battery material, $LiFePO_4$ absorbs and releases energy by the simultaneous extraction and, respectively, insertion of Li^+ ions and electrons. Hence, the power capability of a lithium battery with this or other electrode materials will depend critically on the rate at which the Li^+ ions and electrons can migrate through the electrolyte and composite electrode structure into the active electrode material. Strategies to increase the low rate performance of bulk $LiFePO_4$ have focused on improving electron transport in the bulk[9] or at the surface of the material[10,11], or on reducing the path length over which the electron and Li^+ ion have to move by using nano-sized materials[12,13]. However, recent evidence indicates that Li^+ transport along the surface may be as important as electron transport: whereas $LiFePO_4$ can in principle exchange Li^+ ions with the electrolyte on all surface facets, Li^+ ions can only move into the bulk of the crystal in the [010] direction[7,8,14]. Hence, increasing diffusion across the surface towards the (010) facet should enhance rate capability.

In a departure from previous approaches[9–11], we have created a lithium phosphate coating on the surface of nanoscale $LiFePO_4$ and show that this results in extremely high rate performance. In particular, glassy lithium phosphates are well known to be good, stable Li^+ conductors[15] and can be doped with transition metals to achieve electronic conduction[16–18]. Supplementary Fig. 1 shows a small section of the calculated lithium–iron–phosphorus ternary phase diagram[19] equilibrated with an oxygen potential under reducing conditions, which represents typical synthesis conditions for $LiFePO_4$. Compositions with high phosphorus content on the Li_2O–P_2O_5 binary edge are known to be very good glass formers with high lithium conductivity[20], and nitrogen-doped Li_3PO_4 has been used as a solid-state lithium electrolyte[15]. Typically, the glass-forming ability and lithium conductivity decrease with the presence of Li_2O. These glasses can dissolve a large quantity of transition-metal ions to increase the electronic conductivity[17,21], although such fully amorphous states with high levels of transition metals are usually only obtained by rapid quenching from the liquid state. Hence, the shaded area in the phase diagram represents the optimal coating compositions with good lithium ion conductivity.

Our synthesis strategy has been to create an appropriate off-stoichiometry in the starting materials so that the coating constituents phase-separate from $LiFePO_4$ as it forms during the heat treatment, thereby creating the active storage material and coating in a single process. Here we describe results with an iron:phosphorus deficiency ratio of 2:1 (for example $LiFe_{1-2y}P_{1-y}O_{4-\delta}$, $y = 0.05$), as indicated by arrow A in Supplementary Fig. 1. We note that the more common one-to-one iron:phosphorus deficiency (arrow B in Supplementary Fig. 1, equivalent to lithium excess[22]) creates a mixture of Li_3PO_4 and iron oxides, which are not likely to conduct well under the synthesis conditions used to prepare $LiFePO_4$.

$LiFe_{0.9}P_{0.95}O_{4-\delta}$ was synthesized by ball-milling Li_2CO_3, $FeC_2O_4 \cdot 2H_2O$ and $NH_4H_2PO_4$ in appropriate amounts, heating the mixture at 350 °C for 10 h and then heating at 600 °C for 10 h under argon. X-ray diffraction (Fig. 1a and Supplementary Fig. 2) shows that despite the off-stoichiometric starting mixture, stoichiometric $LiFePO_4$ forms with lattice parameters ($a = 10.3134$ Å, $b = 6.002$ Å and $c = 4.691$ Å) very similar to those reported in the literature[23]. No crystalline Fe_2P can be observed in the X-ray pattern of the material synthesized at 600 °C, but a small amount of Fe_2P is present in the material synthesized at 700 °C (Fig. 1a). However, amorphous FeP or Fe_2P created by the reducing atmosphere cannot be excluded[24]. Mössbauer spectroscopy (Supplementary Fig. 4) indicates that apart from the major $LiFePO_4$ component, around 10% of the Fe is present in some other environment. The isomer shift (0.464 mm s^{-1}) and quadrupole splitting (0.798 mm s^{-1}) of this second component fall in the region of values given in the literature for Fe^{3+} in pyrophosphate (P_2O_7-containing) glasses, although recent work[24] argues that iron monophosphides also give a Mössbauer signal in this range. To distinguish between the two possibilities, as-made material was discharged. The large discharge capacity found (Supplementary Fig. 5) is consistent with the presence of reducible Fe^{3+} rather than FeP in the material. Furthermore, in subsequent charge–discharge cycles we consistently find 15–18 mA h g^{-1} capacity in the 3.2–2.0 V voltage window, in agreement with the ~10% proportion of iron found in the second Mössbauer component. Pyrophosphates are known to have somewhat lower potential than $LiFePO_4$[25].

Particle size as determined by scanning electron microscopy is ~50 nm (Fig. 1b). Transmission electron microscopy (Fig. 1c) shows a poorly crystallized thin layer on the surface. The thickness of this layer varies. Further evidence for the existence of the poorly crystallized layer is provided by X-ray photoelectron spectroscopy, which selectively analyses the surface of a material, and shows two different phosphorus 2*p* chemical states in our material. One state is close to the phosphorus 2*p* binding energy in $LiFePO_4$, but the second component is at higher energy. This is consistent with the presence of the $(P_2O_7)^{4-}$ groups, where phosphorus has higher binding energy than $LiFePO_4$ (Fig. 2 and

[1]Department of Materials Science and Engineering, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139, USA.

©2009 Macmillan Publishers Limited. All rights reserved

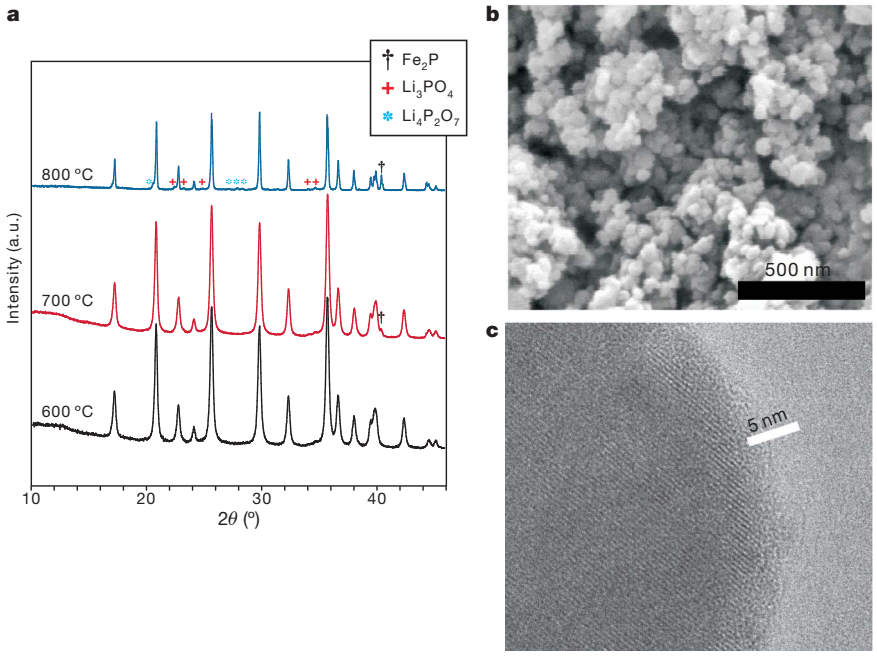


Figure 1 | Characterization of $LiFeP_{0.9}P_{0.95}O_{4-\delta}$ synthesized under argon. **a**, Powder X-ray diffraction patterns (using Cu Kα radiation) for samples synthesized at different temperatures. Fe_2P starts to appear at 700 °C and $Li_4P_2O_7$ starts to crystallize at 800 °C. The lattice parameters for the material synthesized at 600 °C are $a = 10.3134$ Å, $b = 6.002$ Å and $c = 4.691$ Å. θ, diffraction angle; a.u., arbitrary units. **b**, Scanning electron microscopy image showing a particle size of less than 50 nm. **c**, Transmission electron microscopy image of the material showing a poorly crystallized layer less than 5 nm thick on the edge of a particle.

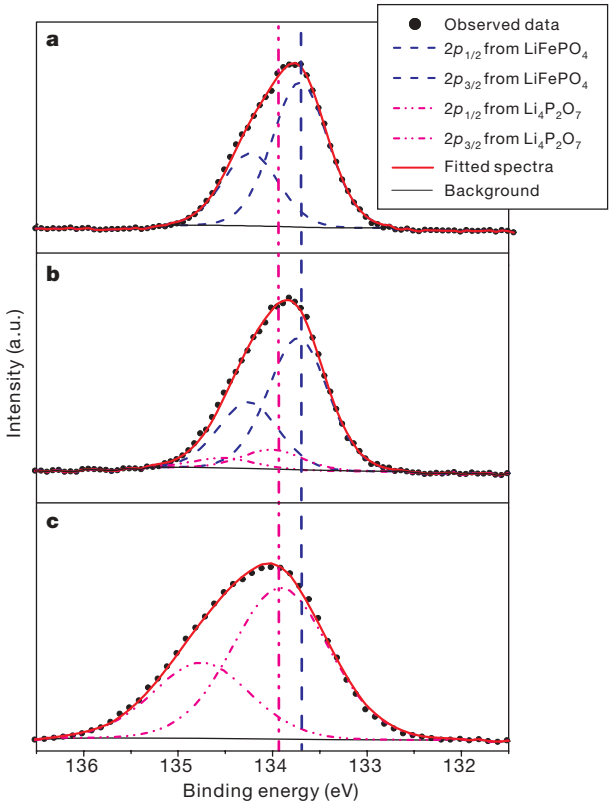


Figure 2 | Phosphorus 2*p* X-ray photoelectron spectra from three different compounds. **a**, $LiFePO_4$; **b**, $LiFe_{0.9}P_{0.95}O_{4-\delta}$; **c**, $Li_4P_2O_7$. The spectra were fitted to the phosphorus 2*p* doublet, $2p_{1/2}$ and $2p_{3/2}$, which is split by 0.84 eV in an integrated intensity ratio of 1:2. The higher energy, lower intensity peak is the $2p_{1/2}$ peak. $LiFe_{0.9}P_{0.95}O_{4-\delta}$ shows two different chemical states for phosphorus. The vertical blue dashed line is the phosphorus $2p_{3/2}$ peak from $LiFePO_4$ and the vertical pink dash–dot line is the phosphorus $2p_{3/2}$ peak from $Li_4P_2O_7$. The data were corrected on the basis of the binding energy of adventitious hydrocarbon, 248.8 eV. The fitted values are in Supplementary Table 1.

Supplementary Table 1)[26]. Elemental analysis of our materials synthesized at 600 °C, performed with scanning transmission electron microscopy (Supplementary Fig. 3), also indicates that the phosphorus:iron ratio on the surface is increased by introducing the off-stoichiometry in the sample. Together, the X-ray photoelectron spectroscopy, Mössbauer spectroscopy and scanning transmission electron microscopy data, as well as the starting stoichiometry, support the idea that the second phase present is a Fe^{3+}-containing $Li_4P_2O_7$-like phase.

Swagelok cells with a metallic lithium anode were assembled to investigate the material. Figure 3a shows the discharge of the material at various rates after a slow charge and hold at 4.3 V to fully charge the material. A rate of *n*C corresponds to a full discharge in $1/n$ h. At a 2C rate, the material still discharges to its estimated theoretical capacity of ~166 mA h g^{-1}. (The theoretical capacity is calculated by assuming a mixture of $LiFePO_4$ and $Li_4P_2O_7$.) Even at the highest rate tested (50C), corresponding to a time of 72 s to fully discharge the capacity, the material achieves about 80% of its theoretical capacity. Capacity retention of the material is superior. As Fig. 3b shows, after 50 full charge–discharge cycles at rates of 20C and 60C, there is no significant capacity loss.

For an electrochemical cell to deliver energy at high rate, all parts of the Li^+–electron path between the anode and cathode active material have to be capable of sustaining this rate. The results in Fig. 3 were obtained when testing a standard cathode preparation using 15 wt% carbon black and 5 wt% polyethylenetetrafluoride as binder. Carbon black is added to facilitate electron transport from the active materials to the current collector. As this preparation has been optimized for materials that have substantially lower rates (~1C) we also tested preparations with up to 65 wt% carbon. Although such high carbon loadings are inappropriate for real batteries, they are useful in establishing the true rate capability of the active material and are commonly used in the testing of high-rate nanomaterials[13,18,27]. The results in Fig. 4 show that extremely high rates can be achieved for the active material: at a 200C rate (corresponding to an 18-s total discharge) more than 100 mA h g^{-1} can still be achieved, and a capacity of 60 mA h g^{-1} is obtained at a 400C rate (9 s to full discharge). Such discharge rates are two orders of magnitude larger than those used in today's lithium ion batteries.

©2009 Macmillan Publishers Limited. All rights reserved

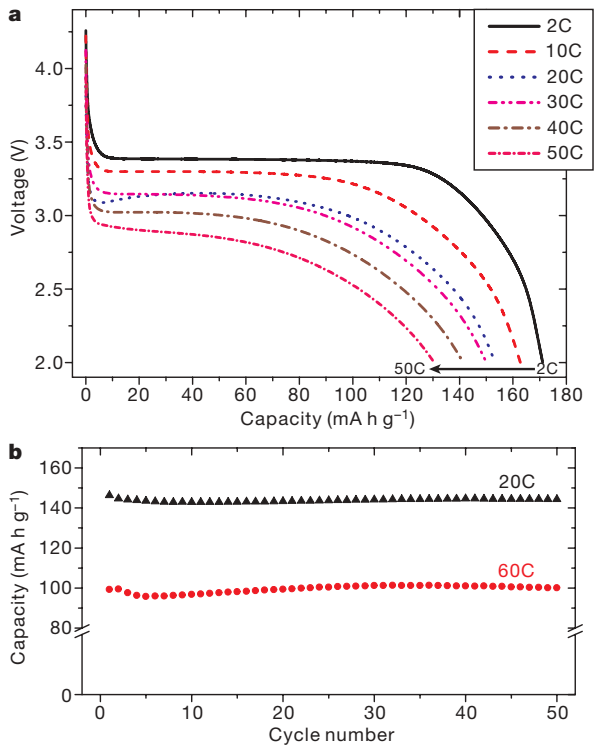


Figure 3 | Discharge rate capability and capacity retention for $LiFe_{0.9}P_{0.95}O_{4-\delta}$ synthesized at 600 °C. **a**, Discharge rate capability after charging at C/5 and holding at 4.3 V until the current reaches C/60. C/*n* denotes the rate at which a full charge or discharge takes *n* hours. The loading density of the electrode is 3.86 mg cm^{-2}. At 2C, the capacity is close to the theoretical value. **b**, Capacity retentions when performing full charge–discharge cycles at constant 20C and 60C current rates for 50 cycles. The loading density of the electrode is 3.60 mg cm^{-2} for the 20C rate and 2.71 mg cm^{-2} for the 60C rate. The voltage window is approximately 2.5–4.3 V. The electrode formulation is active material (80 wt%), carbon (15 wt%) and binder (5 wt%).

Our work provides evidence that extremely high electrochemical discharge rates can be achieved with lithium battery materials. Typical power rates for lithium ion battery materials are in the range of 0.5 to 2 kW kg^{-1}. The specific power we observed for the modified $LiFePO_4$ (170 kW kg^{-1} at a 400C rate and 90 kW kg^{-1} at a 200C rate) is two orders of magnitude higher, but consistent with the very high lithium diffusivity estimated earlier from theoretical calculations[7,8]. Taking the estimated value of the lithium diffusion constant in the [010] direction to be $\sim 10^{-8}$ cm^2 s^{-1}, from ref. 8, we estimate the time for Li^+ to diffuse over 50 nm (approximate particle size) to be [50 nm]^2/[10^{-8} cm^2 s^{-1}], that is, <1 ms. Hence, the limiting factor for charge/discharge is the delivery of Li^+ and electrons to the surface rather than bulk diffusion. This may explain the success of our strategy to facilitate transport across the surface by creating a poorly crystallized layer with high Li^+ mobility. The amorphous nature of the coating removes the anisotropy of the surface properties[28] and enhances delivery of Li^+ to the (010) facet of $LiFePO_4$ where it can be inserted.

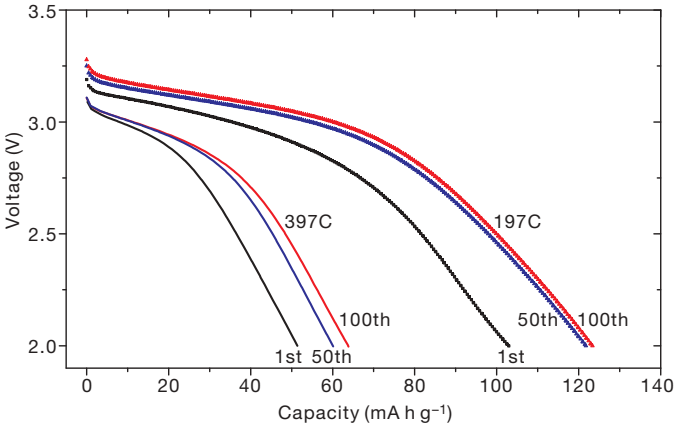


Figure 4 | Discharge capability at very high rate for $LiFe_{0.9}P_{0.95}O_{4-\delta}$ synthesized at 600 °C. Full charge–discharge cycles at constant 197C and 397C current rates without holding the voltage. The loading density of the electrode is 2.96 mg cm^{-2}. The first, fiftieth and hundredth discharges are shown for each rate. The electrode formulation is active material (30 wt%), carbon black (65 wt%) and binder (5 wt%).

It is also plausible that the disordered nature of the coating material modifies the surface potential of lithium to facilitate the adsorption of Li^+ from the electrolyte by providing different lithium sites with a wide range of energies that can be matched to the energy of lithium in the electrolyte. The fact that the particles of our material are nano-sized definitely contributes to its extreme discharge rate capability, but its performance is substantially better than results reported in the literature for particles of similar[9] or smaller size[13], indicating that the coating also enhances the rate capability. To further test that the pyrophosphate coating is responsible for the ultrahigh power rate, we evaluated stoichiometric $LiFePO_4$ synthesized under the same conditions. In this case, the precursors were ball-milled so that grain size after sintering would be the same as for the off-stoichiometric material (Supplementary Fig. 6a). Although the rate capability of this material is good, it is clearly inferior to the off-stoichiometric material (Supplementary Fig. 6b).

Because limited electron transport through the electrode assembly can mask the true rate capability of the material[29,30], our highest rates (200C and 400C) could only be tested with large amounts of carbon, which reduces the volumetric energy density of the electrode. This problem can be addressed by developing electrode structures with good electronic conductivity and percolation while optimizing the volume fraction of the active energy-storing component.

The ability to charge and discharge batteries in a matter of seconds rather than hours may make possible new technological applications and induce lifestyle changes. Such changes may first take place in the use of small devices, where the total amount of energy stored is small. Only 360 W is required to charge a 1 W h cell phone battery in 10 s (at a 360C charging rate). On the other hand, the rate at which very large batteries such as those planned for plug-in hybrid electric vehicles can be charged is likely to be limited by the available power: 180 kW is needed to charge a 15 kW h battery (a typical size estimated for a plug-in hybrid electric vehicle) in 5 min.

Electrode materials with extremely high rate capability will blur the distinction between supercapacitors and batteries. The power density based on the measured volume of the electrode film, including carbon and binder, is around 65 kW l^{-1} in the 400C test. Assuming that the cathode film takes up about 40% of the volume of a complete cell, this will give a power density of $\sim$25 kW l^{-1} at the battery level, which is similar to or higher than the power density in a supercapacitor, yet with a specific energy and energy density one to two orders of magnitude higher. The fact that our material can obtain power densities similar to those of supercapacitors is consistent with there being an exceedingly fast bulk process. For $LiFePO_4$, bulk lithium transport is so fast that the charging is ultimately limited by the surface adsorption and surface transfer, which is also the rate-limiting step in supercapacitors.

METHODS SUMMARY

We synthesized the $LiFe_{1-2y}P_{1-y}O_{4-\delta}$ samples by solid-state reaction using Li_2CO_3 (Fisher Scientific company), $FeC_2O_4 \cdot 2H_2O$ (99.999%, Alfa Aesar) and $NH_4H_2PO_4$ (98%, Alfa Aesar). The mixture in the acetone was ball-milled and heated at 350 °C for 10 h under argon to decompose the carbonate, oxalate and ammonium. The sample was cooled to room temperature, ground and manually

©2009 Macmillan Publishers Limited. All rights reserved

pelletized under 10,000-lb pressure using a disk-shaped mould. After preheating, we heated the pellet at 600 °C for 10 h under argon. Electrodes were made by mixing active material (80 wt%), carbon (carbon black; 15 wt%) and binder (polyethylenetetrafluoride; 5 wt%). The very high rate tests (200C and 400C) were performed on electrodes with 30 wt% active material, 65 wt% carbon and 5 wt% binder. We assembled the cells in an argon-filled glove box and tested them on a Maccor 2200 operating in galvanostatic mode using lithium metal as an anode, non-aqueous electrolyte (1 M $LiPF_6$ in ethylene carbonate:dimethyl carbonate (1:1), Merck) and Celgard 2600 or 2500 as separator in a Swagelok cell. All cells were tested at room temperature.

Received 18 June 2007; accepted 2 February 2009.

1. Conway, B. E. Transition from supercapacitor to battery behavior in electrochemical energy-storage. *J. Electrochem. Soc.* **138,** 1539–1548 (1991).
2. Arico, A. S., Bruce, P., Scrosati, B., Tarascon, J. M. & Van Schalkwijk, W. Nanostructured materials for advanced energy conversion and storage devices. *Nature Mater.* **4,** 366–377 (2005).
3. Amatucci, G. G., Badway, F., Du Pasquier, A. & Zheng, T. An asymmetric hybrid nonaqueous energy storage cell. *J. Electrochem. Soc.* **148,** A930–A939 (2001).
4. Tarascon, J. M. & Armand, M. Issues and challenges facing rechargeable lithium batteries. *Nature* **414,** 359–367 (2001).
5. Reed, J. & Ceder, G. Role of electronic structure in the susceptibility of metastable transition-metal oxide structures to transformation. *Chem. Rev.* **104,** 4513–4533 (2004).
6. Padhi, A. K., Nanjundaswamy, K. S. & Goodenough, J. B. Phospho-olivines as positive-electrode materials for rechargeable lithium batteries. *J. Electrochem. Soc.* **144,** 1188–1194 (1997).
7. Islam, M. S., Driscoll, D. J., Fisher, C. A. J. & Slater, P. R. Atomic-scale investigation of defects, dopants, and lithium transport in the $LiFePO_4$ olivine-type battery material. *Chem. Mater.* **17,** 5085–5092 (2005).
8. Morgan, D., Van der Ven, A. & Ceder, G. Li conductivity in Li_xMPO_4 (M = Mn, Fe, Co, Ni) olivine materials. *Electrochem. Solid State Lett.* **7,** A30–A32 (2004).
9. Chung, S. Y., Bloking, J. T. & Chiang, Y. M. Electronically conductive phospho-olivines as lithium storage electrodes. *Nature Mater.* **1,** 123–128 (2002).
10. Ravet, N. C. Y., Magnan, J. F., Besner, S., Gauthier, M. & Armand, M. Electroactivity of natural and synthetic triphylite. *J. Power Sources* **97–8,** 503–507 (2001).
11. Herle, P. S., Ellis, B., Coombs, N. & Nazar, L. F. Nano-network electronic conduction in iron and nickel olivine phosphates. *Nature Mater.* **3,** 147–152 (2004).
12. Delacourt, C., Poizot, P., Levasseur, S. & Masquelier, C. Size effects on carbon-free $LiFePO_4$ powders. *Electrochem. Solid State Lett.* **9,** A352–A355 (2006).
13. Kim, D. H. & Kim, J. Synthesis of $LiFePO_4$ nanoparticles in polyol medium and their electrochemical properties. *Electrochem. Solid State Lett.* **9,** A439–A442 (2006).
14. Chen, G. Y., Song, X. Y. & Richardson, T. J. Electron microscopy study of the $LiFePO_4$ to $FePO_4$ phase transition. *Electrochem. Solid State Lett.* **9,** A295–A298 (2006).
15. Wang, B., Kwak, B. S., Sales, B. C. & Bates, J. B. Ionic conductivities and structure of lithium phosphorus oxynitride glasses. *J. Non-Cryst. Solids* **183,** 297–306 (1995).
16. Sayer, M. & Mansingh, A. Transport properties of semiconducting phosphate glasses. *Phys. Rev. B* **6,** 4629–4643 (1972).
17. Mogus-Milankovic, A., Santic, A., Karbulut, M. & Day, D. E. Study of electrical properties of MoO_3-Fe_2O_3-P_2O_5 and SrO-Fe_2O_3-P_2O_5 glasses by impedance spectroscopy. II. *J. Non-Cryst. Solids* **330,** 128–141 (2003).
18. Zhou, H. S., Li, D. L., Hibino, M. & Honma, I. A self-ordered, crystalline-glass, mesoporous nanocomposite for use as a lithium-based storage device with both high power and high energy densities. *Angew. Chem. Int. Edn Engl.* **44,** 797–802 (2005).
19. Ong, S. P., Wang, L., Kang, B. & Ceder, G. Li-Fe-P-O_2 phase diagram from first principles calculations. *Chem. Mater.* **20,** 1798–1807 (2008).
20. Martin, S. W. Ionic-conduction in phosphate-glasses. *J. Am. Ceram. Soc.* **74,** 1767–1784 (1991).
21. Sobha, K. C. & Rao, K. J. Investigation of phosphate glasses with the general formula $A_xB_yP_3O_{12}$ where A=Li, Na or K and B=Fe, Ga, Ti, Ge, V or Nb. *J. Non-Cryst. Solids* **201,** 52–65 (1996).
22. Kim, D.-K. *et al.* Effect of synthesis conditions on the properties of $LiFePO_4$ for secondary lithium batteries. *J. Power Sources* **159,** 237–240 (2006).
23. Ellis, B. *et al.* Nanostructured materials for lithium-ion batteries: Surface conductivity vs. bulk ion/electron transport. *Faraday Discuss.* **134,** 119–141 (2007).
24. Rho, Y. H., Nazar, L. F., Perry, L. & Ryan, D. Surface chemistry of $LiFePO_4$ studied by Mossbauer and X-ray photoelectron spectroscopy and its effect on electrochemical properties. *J. Electrochem. Soc.* **154,** A283–A289 (2007).
25. Padhi, A. K., Nanjundaswamy, K. S., Masquelier, C., Okada, S. & Goodenough, J. B. Effect of structure on the Fe^{3+}/Fe^{2+} redox couple in iron phosphates. *J. Electrochem. Soc.* **144,** 1609–1613 (1997).
26. Morgan, W. E., Stec, W. J. & Vanwazer, J. R. Inner-orbital photoelectron spectroscopy of alkali-metal halides, perchlorates, phosphates, and pyrophosphates. *J. Am. Chem. Soc.* **95,** 751–755 (1973).
27. Okubo, M. *et al.* Nanosize effect on high-rate Li-ion intercalation in $LiCoO_2$ electrode. *J. Am. Chem. Soc.* **129,** 7444–7452 (2007).
28. Wang, L., Zhou, F., Meng, Y. S. & Ceder, G. First-principles study of surface properties of $LiFePO_4$: Surface energy, structure, Wulff shape, and surface redox potential. *Phys. Rev. B* **76,** 165435 (2007).
29. Dominko, R., Gaberscek, M., Bele, A., Mihailovic, D. & Jamnik, J. Carbon nanocoatings on active materials for Li-ion batteries. *J. Eur. Ceram. Soc.* **27,** 909–913 (2007).
30. Gaberscek, M., Dominko, R., Bele, M., Remskar, M. & Jamnik, J. Mass and charge transport in hierarchically organized storage materials. Example: Porous active materials with nanocoated walls of pores. *Solid State Ionics* **177,** 3015–3022 (2006).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements B.K. thanks D. S. Yun and Y. Zhang for the help with transmission electron microscope measurements and K. Kang and Y. S. Meng for experimental help and discussions. Support from the US National Science Foundation through the Materials Research Science and Engineering Centers programme and the Batteries for Advanced Transportation Program of the US Department of Energy is gratefully acknowledged.

Author Contributions B.K. performed the experiments and G.C. supervised and analysed the work.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to G.C. (gceder@mit.edu).

©2009 Macmillan Publishers Limited. All rights reserved

LETTERS

Uplift of the Longmen Shan and Tibetan plateau, and the 2008 Wenchuan ($M = 7.9$) earthquake

Judith Hubbard[1] & John H. Shaw[1]

The Longmen Shan mountain range, site of the devastating 12 May 2008 Wenchuan ($M = 7.9$) earthquake, defines the eastern margin of the Himalayan orogen and exhibits greater topographic relief than anywhere else in the Tibetan plateau. However, before the earthquake, geodetic and geologic surveys measured little shortening across the range front[1–3], inspiring a vigorous debate about the process by which the topography of the mountain belt is produced and maintained. Two endmember models have been proposed: (1) brittle crustal thickening, in which thrust faults with large amounts of slip that are rooted in the lithosphere cause uplift[4], and (2) crustal flow, in which low-viscosity material in the lower crust extrudes outward from the Tibetan plateau and inflates the crust north and east of the Himalayas[5–7]. Here we use balanced geologic cross-sections to show that crustal shortening, structural relief, and topography are strongly correlated in the range front. This suggests that crustal shortening is a primary driver for uplift and topography of the Longmen Shan on the flanks of the plateau. The 2008 Wenchuan ($M = 7.9$) earthquake, which ruptured a large thrust fault along the range front causing tens of thousands of fatalities and widespread damage, is an active manifestation of this shortening process.

India began colliding with Eurasia over 50 million years ago, progressively accommodating at least 1,400 km of north–south shortening[8]. This plate collision uplifted the highest mountains in the world and also a vast flat region to the north known as the Tibetan plateau. This plateau has an average elevation of over 5,000 m, maintained in part by a 65–75-km-thick crust with a complex and heterogeneous three-dimensional structure inferred from tomographic studies.[9–11]

On the eastern side of the Tibetan plateau, the Longmen Shan rise 6,000 m above the Sichuan basin, exhibiting greater relief than anywhere else on the plateau (Fig. 1). Exhumation rates suggest that the relief has formed since the Miocene epoch[12]. However, geodetic measurements[1–3] and geomorphic observations[13] generally agree that east–west shortening across the range and within the basin is very limited (<3 mm yr^{-1}). Moreover, the Longmen Shan lacks a typical foreland basin, a depression caused by plate loading that typically occurs in front of fold-and-thrust belts[14,15]. As a result, basic and controversial questions persist about how this relief was created and is sustained, fuelling debate about how the broader Tibetan plateau was formed. Are the mountains formed by crustal faulting and shortening[4], or are they maintained by dynamic pressure from lower-crustal flow[5–7]? The second hypothesis argues that lower-crustal material flowing outward from the centre of the plateau is buttressed by the old, strong lithosphere that underlies the Sichuan basin, pushing up on the crust above and maintaining steep topography through dynamic pressure. This model requires no substantial horizontal shortening at the surface, because the upper crust is largely uplifted, but not laterally displaced, by the lower-crustal motion[16]. (Fig. 2)

An important distinction between these two models is that brittle crustal thickening requires horizontal shortening to be related to vertical uplift in the Longmen Shan, whereas lower-crustal flow suggests that the two are largely decoupled[16,17]. Here we aim to quantify past horizontal shortening on the eastern border of the Longmen Shan to determine whether the gradient in horizontal shortening reflects the current topography. Measures of horizontal shortening are derived from structural interpretation of industry seismic reflection profiles and surface geology in the Sichuan basin and the foothills of the Longmen Shan.

Substantial amounts of industry seismic reflection data have been acquired since the 1980s in the Sichuan basin and eastern Longmen Shan to identify petroleum traps; these data offer clear images of faults and related folds in the upper crust (see Supplementary Figs 1 and 2). The lines are supported by a number of wells providing chrono-stratigraphic control on the subsurface geology. We use quantitative fault-related folding techniques to interpret these data

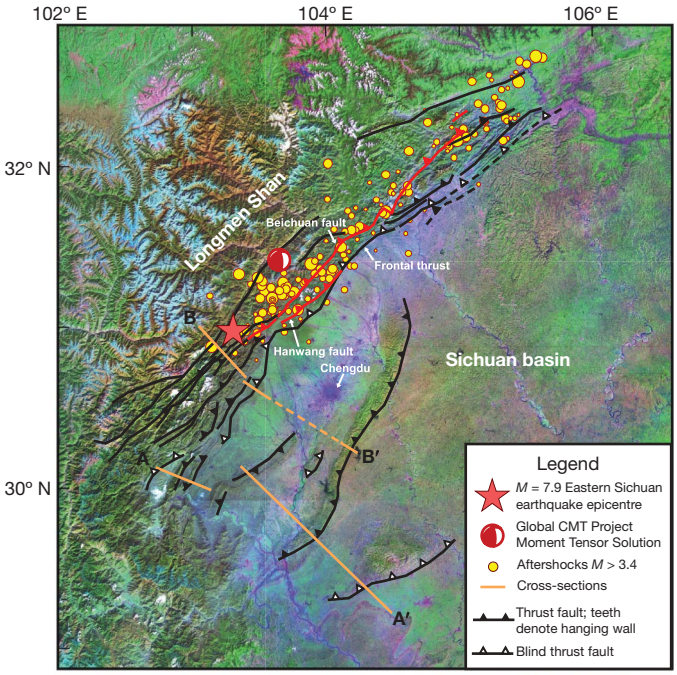


Figure 1 | Map of the Longmen Shan and western Sichuan basin, showing the epicentre and focal mechanism of the 2008 Wenchuan ($M = 7.9$) earthquake, aftershocks, and major faults. A–A′ and B–B′ are cross-sections based on seismic reflection profiles used to characterize crustal deformation in this study. Emergent and blind thrusts are distinguished; faults in red are the Beichuan and Hanwang faults, which ruptured in the 2008 Wenchuan earthquake.

[1]Department of Earth and Planetary Sciences, Harvard University, 20 Oxford Street, Cambridge, Massachusetts 02138, USA.

©2009 Macmillan Publishers Limited. All rights reserved

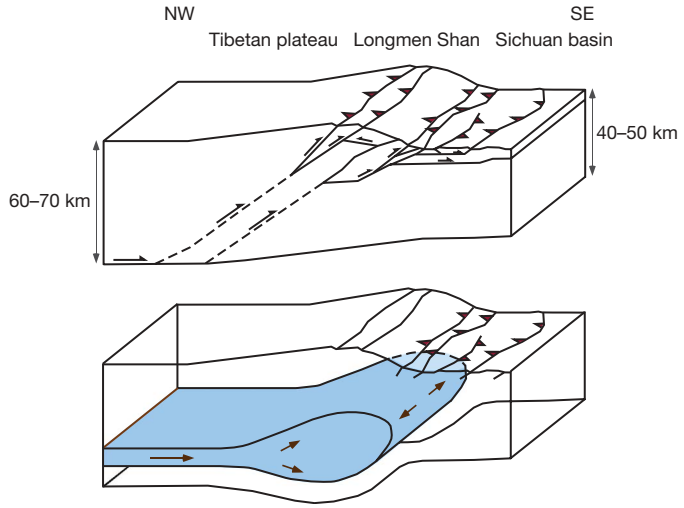


Figure 2 | Alternative conceptual models for the uplift of the Tibetan plateau and development of the Longmen Shan range front. Top panel, uplift is produced by thrust faulting and crustal shortening[4]. Bottom panel, uplift is produced by inflation of the ductile lower crust[5–7].

and construct balanced, retrodeformable cross-sections[18,19]. These balanced sections provide direct measures of shortening and uplift, which we examine in conjunction with topography across the range front.

Deformation in the southwestern Sichuan basin can be separated into six northeast–southwest-striking domains, visible in both the seismic data and satellite images of the surface (Fig. 1). From east to west, these are (1) a set of broad folds, hereafter referred to as the Zigong folds after a city in their vicinity, (2) the Weiyuan anticline, (3) the Longqan anticline, (4) the Emei anticline, (5) the Xiong Po anticline, and (6) the Longmenshan foothills (Fig. 3b). In addition, the strata exhibit a regional west dip of two to three degrees towards the Longmen Shan west of the Weiyuan structure. This dip may be related to loading of the crust by the mountain belt. However, it lacks the typical expression of a foreland basin, perhaps due to the lack of synorogenic sediments, which are not present in the basin.

The six major structural trends are formed by fault-bend or fault-propagation folds rooted in detachments that transfer shortening from the mountain belt towards the east into the Sichuan basin. From the seismic reflection data, we identify three detachment levels, which define the base of the thrust sheets: a shallow detachment 3–5 km below the surface, an intermediate detachment 5–7 km below the surface, and a deep detachment at least 15 km deep on the northwestern side of the lines, shallowing to no less than 6 km to the southeast. Each detachment feeds slip onto thrust ramps that produce upper crustal shortening and uplift the range front of the Longmen Shan. The amounts of slip interpreted on the thrust ramps and detachments are constrained by the offsets of hanging-wall and footwall cut-offs of stratigraphic horizons, as well as the shapes of associated fault-related folds.

We focus our discussion of shortening and structural relief on a regional structural transect (lines A–A′ in Fig. 1) that represents the general style of deformation along the range front. The total shortening measured by line length on the six structures totals ~25.8 km, with most taken up across the Longmenshan foothills (13.3 km) and the Xiong Po backthrust (~7.4 km), the structures closest to the range front. The Emei, Longqan, Weiyuan and Zigong structures account for 1.0, 1.1, 1.1 and 1.9 km of shortening, respectively.

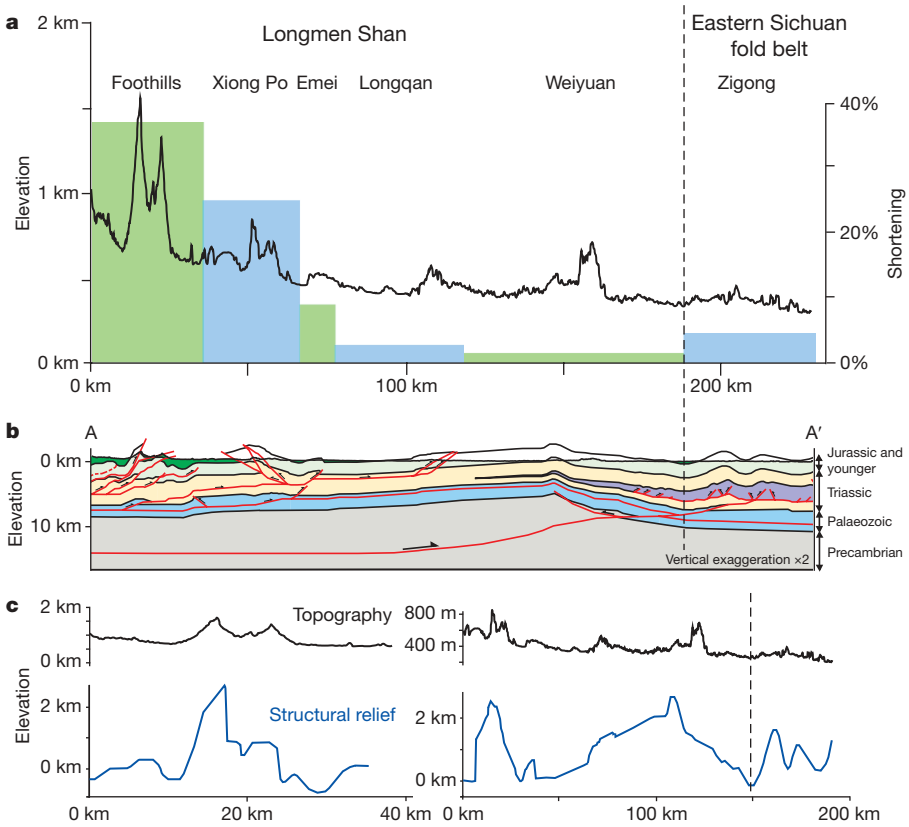


Figure 3 | Comparison of crustal shortening and topography along line A–A′. See Table 1. **a**, Percentage shortening is shown as colour bars (alternating green and blue) plotted over topography; we note the dramatic increase in shortening towards the range front. **b**, Geologic cross-section. **c**, Topography (black) and structural relief (blue) from structural interpretation. Vertical dashed lines separate the Zigong folds from the rest of the line (see text).

©2009 Macmillan Publishers Limited. All rights reserved

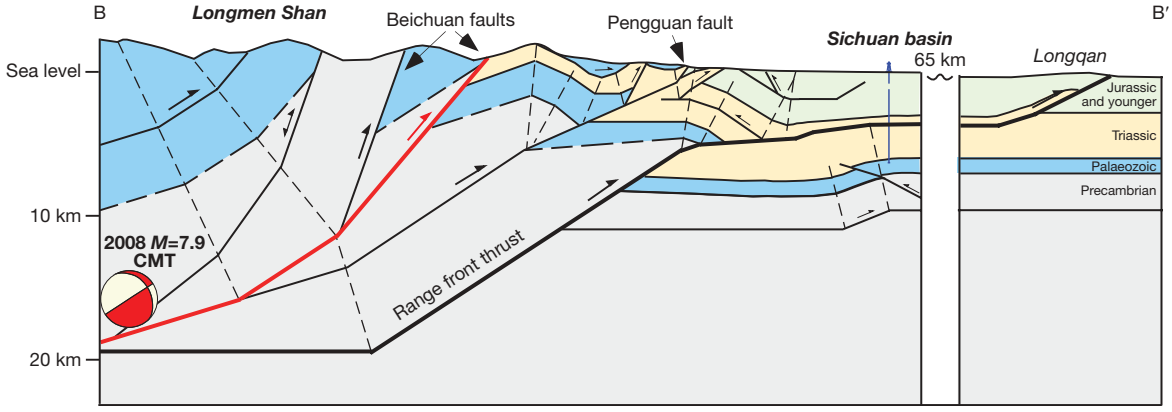


Figure 4 | Geologic cross-section B–B′ across the location of the 2008 Wenchuan ($M = 7.9$) earthquake. On the basis of the depth and preferred nodal plane of the main shock, the earthquake and its aftershocks appear to have ruptured one or more imbricated thrust and reverse faults in the range front, including the Beichuan faults, of which there are reports of surface rupture. 'CMT' indicates the Global Centroid Moment Tensor (CMT) Project (http://www.globalcmt.org/) Moment Tensor Solution. The Pengguan fault lies south of the Hanwang fault but did not rupture in the earthquake.

To determine the gradient of shortening, we calculate the percentage shortening for each structural trend. Crustal shortening is defined as the amount of northwest-to-southeast contraction that has occurred across each of the six major structural domains as the result of folding and faulting. The shortening measurements vary from about 2% up to 40%, and dramatically increase upon approaching and entering the Longmen Shan. This westward increase in crustal shortening directly corresponds with the rapidly increasing elevations of the range front, suggesting that brittle shortening is closely linked to producing and maintaining the topography of the mountain belt (Table 1 and Fig. 3a).

In addition, we calculate profiles of structural relief along lines A–A′ and compare them with topography (Fig. 3c). Structural relief records the total uplift produced by each fault and fold trend in the basin, and thus is considered to be another measure of upper crustal deformation. We quantify structural relief using the deformed shape of the youngest, pre-tectonic Jurassic stratigraphic horizon that is defined by continuous reflections in the industry seismic data. The patterns of structural relief correspond closely with the present-day topography observed along the range front. The magnitudes and gradients in topographic relief are clearly less than the structural relief due to erosion. Nevertheless, the positions of the maxima, minima and gradients in structural relief and topography closely correlate, implying that present-day topography is produced by uplift along faults and folds in the upper crust. We note that the Zigong folds, which represent the westernmost extent of the largely inactive Eastern Sichuan fold belt rather than deformation associated with the Longmen Shan, do not show this correlation.

In summary, our quantitative interpretation of structures in the Sichuan basin and foothills of the Longmen Shan indicates that a large increase in shortening occurs upon approaching the range front. Moreover, structural relief produced by folds and faults is also closely correlated with, and sufficient in magnitude to explain, the topography of the range front. This implies that upper-crustal deformation, decoupled from the lower crust by a series of detachments, is the primary mechanism for generating uplift and topography in the foothills of the Longmen Shan. In the crustal shortening versus lower-crustal-flow debate, it is the range front's reported absence of active shortening that underpins the lower-crustal-flow hypothesis, and some models proposed for this lower-crustal-flow process explicitly prescribe no correlation between upper-crustal shortening and topography[16].

Thus, we interpret our results to obviate the need for lower-crustal flow and inflation to produce and maintain the Longmen Shan range front, and argue that if lower-crustal flow occurs in the interior of the mountain belt, it must somehow be directly coupled to upper-crustal deformation in the range front. Alternatively, further deepening of the active thrust systems and shortening west of our study area, which has been inferred in recent geodetic studies[14] and documented in geologic mapping[20], may account for uplift of the plateau in the absence of significant lower-crustal flow.

The Wenchuan ($M = 7.9$) earthquake is a direct manifestation of the active crustal shortening that we suggest produced and supports the Longmen Shan range front. This devastating earthquake occurred on a series of large, crustal thrust faults that form the range front and accommodate crustal shortening immediately west of the area we considered (Fig. 4). From the distribution of large aftershocks and reported surface ruptures, it appears that this earthquake and its aftershocks caused slip on two major faults within this imbricated thrust stack that ramp upward to the surface, as well as others that transfer slip over geologic timescales into the basin on the detachment surfaces documented here.

Thus, the Wenchuan ($M = 7.9$) earthquake documents the role of active crustal shortening in developing and supporting the Longmen Shan range front and the Tibetan plateau, consistent with the results of our analysis that relate crustal shortening to the development of range front topography. This earthquake also clearly illustrates the potential for large and very destructive events on similar active thrust faults that have been observed in other populated regions around the world.

Table 1 | Shortening measurements for each structure along line A–A′

Structure	Shortening (m)	Distance (km)	Shortening (%)
Foothills	13.3	36	37
Xiong Po	7.4	30	25
Emei	1.0	11	9
Longqan	1.1	41	2.7
Weiyuan	1.1	69	1.6
Zigong	1.9	41	4.6

The percentage shortening was calculated by dividing shortening by the horizontal width of each structural belt.

Received 11 June 2008; accepted 27 January 2009.

1. Shen, Z. K., Lu, J. N., Wang, M. & Burgmann, R. Contemporary crustal deformation around the southeast borderland of the Tibetan Plateau. *J. Geophys. Res.* **110,** 11409, doi:10.1029/2004JB003421 (2005).
2. Meade, B. J. Present-day kinematics at the India-Asia collision zone. *Geology* **35,** 81–84 (2007).
3. Chen, Z. *et al.* Global Positioning System measurements from eastern Tibet and their implications for India/Eurasia intercontinental deformation. *J. Geophys. Res.* **105,** 16215–16227 (2000).
4. Tapponnier, P. *et al.* Geology—Oblique stepwise rise and growth of the Tibet plateau. *Science* **294,** 1671–1677 (2001).
5. Royden, L. H. *et al.* Surface deformation and lower crustal flow in eastern Tibet. *Science* **276,** 788–790 (1997).
6. Burchfiel, B. C. 2003 presidential address: New technology, new geological challenges. *GSA Today* **14,** 4–9 (2004).
7. Bird, P. Lateral extrusion of lower crust from under high topography, in the isostatic limit. *J. Geophys. Res.* **96,** 10275–10286 (1991).

©2009 Macmillan Publishers Limited. All rights reserved

8. Yin, A. & Harrison, T. M. Geologic evolution of the Himalayan-Tibetan orogen. *Annu. Rev. Earth Planet. Sci.* **28,** 211–280 (2000).
9. Rapine, R., Tilmann, F., West, M., Ni, J. & Rodgers, A. Crustal structure of northern and southern Tibet from surface wave dispersion analysis. *J. Geophys. Res.* **108,** B22120, doi:10.1029/2001JB000445 (2003).
10. Xu, L., Rondenay, S. & van der Hilst, R. D. Structure of the crust beneath the southeastern Tibetan Plateau from teleseismic receiver functions. *Phys. Earth Planet. Inter.* **165,** 176–193 (2007).
11. Yao, H., Behein, C. & van der Hilst, R. D. Surface wave array tomography in SE Tibet from ambient seismic noise and two-station analysis. II. Crustal and upper-mantle structure. *Geophys. J. Int.* **173,** 205–219 (2008).
12. Kirby, E. *et al.* Late Cenozoic evolution of the eastern margin of the Tibetan Plateau: Inferences from $^{40}Ar/^{39}Ar$ and (U-Th)/He thermochronology. *Tectonics* **21** (1), 1001, doi:10.1029/2000TC001246 (2002).
13. Densmore, A. L. *et al.* Active tectonics of the Beichuan and Pengguan faults at the eastern margin of the Tibetan Plateau. *Tectonics* **26,** TC4005, doi:10.1029/2006TC001987 (2007).
14. Burchfiel, B. C. *et al.* A geological and geophysical context for the Wenchuan earthquake of 12 May 2008, Sichuan, People's Republic of China. *GSA Today* **18,** 4–11 (2008).
15. Burchfiel, B. C. *et al.* Tectonics of the Longmen Shan and adjacent regions, Central China. *Int. Geol. Rev.* **37,** 661–735 (1995).
16. Clark, M. K., Bush, J. W. M. & Royden, L. H. Dynamic topography produced by lower crustal flow against rheological strength heterogeneities bordering the Tibetan Plateau. *Geophys. J. Int.* **162,** 575–590 (2005).
17. Medvedev, S. & Beaumont, C. in *Channel Flow, Ductile Extrusion and Exhumation in Continental Collision Zones* (eds Law, R. D., Searle, M. P. & Godin, L.) 147–164 (Geological Society, London, 2006).
18. Shaw, J. H., Connors, C. & Suppe, J. *Seismic Interpretation of Contractional Fault-related Folds: An AAPG Seismic Atlas* 1-58 (AAPG, 2005).
19. Suppe, J. Geometry and kinematics of fault-bend folding. *Am. J. Sci.* **283,** 684–721 (1983).
20. Chen, S. F. & Wilson, C. J. L. Emplacement of the Longmen Shan Thrust-Nappe Belt along the eastern margin of the Tibetan Plateau. *J. Struct. Geol.* **18,** 413–430 (1996).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank the reviewers, E. Kirby and J. Dolan, for their comments and suggestions. Texaco and the China National Petroleum Corporation provided the seismic data used in this project.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to J.H. (jhubbard@fas.harvard.edu).

©2009 **Macmillan Publishers Limited. All rights reserved**

LETTERS

Age of Zhoukoudian *Homo erectus* determined with $^{26}Al/^{10}Be$ burial dating

Guanjun Shen[1], Xing Gao[2], Bin Gao[1] & Darryl E. Granger[3]

The age of Zhoukoudian *Homo erectus*, commonly known as 'Peking Man', has long been pursued, but has remained problematic owing to the lack of suitable dating methods[1–7]. Here we report cosmogenic $^{26}Al/^{10}Be$ burial dating of quartz sediments and artefacts from the lower strata of Locality 1 in the southwestern suburb of Beijing, China, where early representatives of Zhoukoudian *Homo erectus* were discovered. This study marks the first radioisotopic dating of any early hominin site in China beyond the range of mass spectrometric U-series dating. The weighted mean of six meaningful age measurements, 0.77 ± 0.08 million years (Myr, mean ± s.e.m.), provides the best age estimate for lower cultural layers 7–10. Together with previously reported U-series dating of speleothem calcite[3] and palaeomagnetic stratigraphy[4], as well as sedimentological considerations[8,9], these layers may be further correlated to S6–S7 in Chinese loess stratigraphy or marine isotope stages (MIS) 17–19, in the range of ~0.68 to 0.78 Myr ago. These ages are substantially older than previously supposed and may imply early hominin's presence at the site in northern China through a relatively mild glacial period corresponding to MIS 18.

With an inventory of 6 fairly complete hominin crania and bones representing at least 40 individuals, 98 species of non-hominin mammalian fossils and tens of thousands of stone artefacts, the cave site of Zhoukoudian Locality 1 has remained the largest single source of *Homo erectus* and is one of the most important Palaeolithic sites in the world[1]. The site is a sedimentary infill within a vertical karstic fissure. Its ~40-m-thick depositional sequence can be divided into 17 layers[10]. The lowermost layers 11–17 are fluvial, layers 6–10 are breakdown breccia from the cave walls and ceiling interbedded with silt and sand, layer 5 is travertine, and the uppermost layers 1–4 are silt and travertine with minor breakdown that accumulated after collapse of the ceiling[11]. Stone artefacts and hominin fossils have been recovered from layers 1–10, with most from a lower level in layers 8/9 and an upper level in layers 3–4 (ref. 1). Mammalian fossils are found in layers 1–13, with some primitive carnivores disappearing above layer 5 (ref. 12).

As part of a multidisciplinary study initiated in the late 1970s[1], dating was carried out at several Chinese institutions using a variety of techniques. The following age sequence was proposed: ~700 thousand years (kyr) for the lowest fossiliferous layer 13, based mainly on palaeomagnetic stratigraphy[4]; ~500 kyr for the lowest hominin-fossil-bearing layer 10, based on fission-track dating of sphene grains; and ~230 kyr for the uppermost layers 1–3, based on $^{230}Th/^{234}U$ dating of fossil materials[2]. These age assignments were generally supported by later $^{231}Pa/^{235}U$ (ref. 5), fission-track[6] and electron spin resonance dating[7]. A time range of ~230 to 500 kyr ago for the hominin-fossil-bearing layers has been widely accepted by palaeoanthropologists, although with a few critical comments[13].

In contrast, much older ages were determined using mass spectrometric U-series dating of intercalated pure and dense calcite samples, known to be a more reliable chronometer[14,15]. A date of 400 ± 8 kyr ago was proposed for an upper horizon of layers 1/2, ~500 kyr ago for the upper part of layer 5, and ≥600 kyr ago for the middle and lower parts of layer 5 (Fig. 1, ref. 3).

The suggestion that Zhoukoudian *H. erectus* is substantially older than previously estimated remains to be validated by independent checks. However, numerical dating beyond the upper limit of mass spectrometric U-series dating, ~600 kyr ago, is difficult in China because the lack of contemporaneous volcanic activity nearly precludes the application of $^{40}Ar/^{39}Ar$ dating. Here we use burial dating with cosmogenic ^{26}Al and ^{10}Be in quartz[16–19], which is often suitable for allochthonous cave sediments such as those at Locality 1. This method is based on radioactive decay of ^{26}Al ($t_{1/2} = 717 \pm 17$ kyr)[17] and ^{10}Be ($t_{1/2} = 1.36 \pm 0.07$ Myr)[20]. These two nuclides are produced with a known $^{26}Al/^{10}Be$ atomic ratio of 6.8:1 in quartz exposed to secondary cosmic radiation near the ground surface. Their initial concentrations depend on the mineral's exposure time, which in turn is controlled by the erosion rate of the host rock. If quartz grains from the surface are deeply buried, for example by deposition in a cave, then the production of cosmogenic nuclides nearly stops. Because ^{26}Al decays faster than ^{10}Be, the $^{26}Al/^{10}Be$ ratio decreases exponentially with an apparent half-life of 1.52 Myr. This offers a means for dating quartz burial up to ~3–5 Myr ago[16].

Burial dating was first applied to quartz gravels in caves for deriving river incision rates[17]. It was later applied to hominin sites at Sterkfontein in South Africa[18] and Sima del Elefante at Atapuerca in Spain[19]. The strengths of this method are its radiometric basis and its independence from other dating methods. However, it must be recognized that cave sediments can have complex stratigraphy, particularly in vadose fills. If fossils are mixed with quartz sediments with a prior burial history, the resulting age will be erroneously old.

Six quartz-bearing sand samples were collected, two of them (ZKD-12 and ZKD-13) from fluvial deposits in layers 12 and 13 and the other four (ZKD-6, ZKD-7, ZKD-8/9 and ZKD-10) from quartz-rich lenses or sublayers in layers 6, 7, 8/9 and 10, respectively. Moreover, four quartzite artefacts that directly indicate hominin presence were analysed from collections made in the 1930s from layers 8/9. The ^{26}Al and ^{10}Be concentrations and corresponding burial ages are presented in Table 1.

Three of the four quartzite artefacts yield results consistent within 1σ, with an error-weighted mean age of 0.72 ± 0.13 Myr. The fourth artefact (ST-3) gives an aberrant result of 1.66 ± 0.21 Myr. This particular sample could have been taken from an older cave fill or terrace before manufacture. Among the sediment samples, those from layers 7, 8/9 and 10 yield consistent results, with an error-weighted mean age of 0.81 ± 0.11 Myr. This is slightly older than, but within error of, the weighted mean of the results from the three artefacts, indicating that some sand might have entered the cave with a previous burial signal.

[1]College of Geographical Sciences, Nanjing Normal University, Nanjing, Jiangsu 210046, China. [2]Institute of Vertebrate Paleontology and Paleoanthropology, Academia Sinica, Beijing 100044, China. [3]Department of Earth and Atmospheric Sciences, Purdue University, West Lafayette, Indiana 47907-2051, USA.

©2009 Macmillan Publishers Limited. All rights reserved

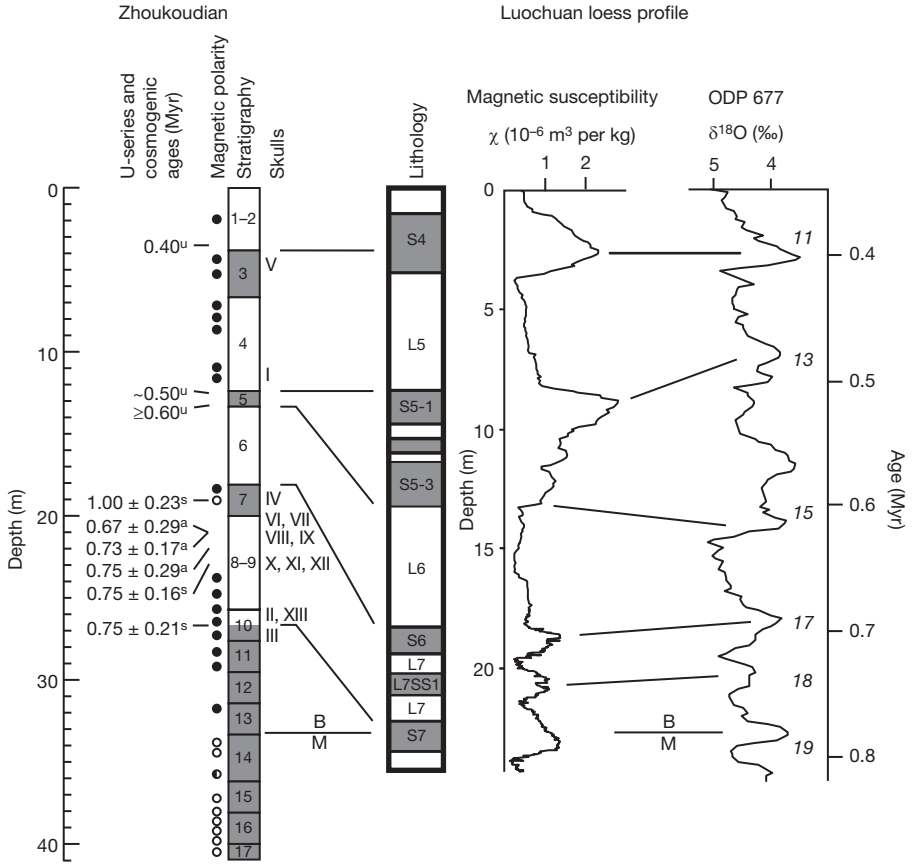


Figure 1 | Correlation of Zhoukoudian Locality 1 with loess stratigraphy and marine isotope record from Ocean Drilling Program (ODP) core 677 (ref. 21). The depositional sequence is constrained in time by the Brunhes/Matuyama (B/M) palaeomagnetic boundary[4], U-series ages of flowstones[3], and cosmogenic burial ages reported here. Shading indicates primarily waterlain layers for Locality 1 stratigraphy[11] and palaeosols for the Luochuan loess profile[22]. Roman numerals denote the locations of hominin skulls[1,11]. Superscripts on ages indicate dating method (a, cosmogenic nuclides in artefacts; s, cosmogenic nuclides in sediment; u, U-series). Magnetostratigraphic data[4] are indicated by closed, open and half-closed circles for normal, reverse and uncertain polarity, respectively. Luochuan loess magnetic susceptibility data (unpublished) are provided courtesy of Dr Liping Zhou. Marine isotope stages are shown in italics.

ZKD-6 gives an aberrant result, 2.78 ± 0.51 Myr. This sample may possibly date to an earlier phase of cave formation, as it was collected from a thin sandy layer that is adhered to the north wall and is now out of stratigraphic contact with the main cross-section. The two samples from the basal fluvial sediments do not yield statistically meaningful results. Their inherited cosmogenic nuclide concentrations are quite low due to rapid erosion in their source area, leading to large uncertainty. Taken together, we consider the weighted mean of the six meaningful measurements, 0.77 ± 0.08 Myr, to best represent the age for layers 7–10. This is consistent with both previous U-series[3] and palaeomagnetic[4] data.

Table 1 | Cosmogenic nuclide concentrations and burial ages

Sample	[^{26}Al] (10^6 at g^{-1})	[^{10}Be] (10^6 at g^{-1})	^{26}Al/^{10}Be	Burial age*(Myr)
Sediment				
ZKD-6	0.073 ± 0.018	0.040 ± 0.004	1.82 ± 0.49	2.78 ± 0.51 (0.54)
ZKD-7-2	0.550 ± 0.053	0.132 ± 0.009	4.17 ± 0.49	1.00 ± 0.23 (0.24)†
ZKD-8/9	1.252 ± 0.095	0.273 ± 0.008	4.58 ± 0.38	0.75 ± 0.16 (0.17)†
ZKD-10-2	0.568 ± 0.052	0.120 ± 0.006	4.72 ± 0.50	0.75 ± 0.21 (0.22)†
ZKD-12	0.105 ± 0.030	0.021 ± 0.006	5.10 ± 2.01	0.62 ± 0.74 (0.74)
ZKD-13	0.106 ± 0.028	0.018 ± 0.005	5.89 ± 2.35	0.31 ± 0.74 (0.74)
Artefacts (8/9)				
ST-1	0.199 ± 0.027	0.040 ± 0.002	4.95 ± 0.72	0.67 ± 0.29 (0.29)†
ST-2	0.476 ± 0.037	0.100 ± 0.003	4.77 ± 0.39	0.73 ± 0.17 (0.17)†
ST-3	0.371 ± 0.039	0.122 ± 0.003	3.04 ± 0.33	1.66 ± 0.21 (0.24)
ST-4	0.568 ± 0.083	0.120 ± 0.005	4.72 ± 0.71	0.75 ± 0.29 (0.30)†

* Uncertainties (±1σ) are expressed in two ways: the first includes analytical uncertainty only, and should be used when comparing burial ages; the second uncertainty listed in parentheses also includes systematic errors in half-lives, and should be used when comparing against ages from other radiometric dating methods.
† Ages included in the weighted mean for layers 7–10.

The above age assignment may be refined in the context of the cave environment. Several previous studies have correlated sedimentary packages at Zhoukoudian Locality 1 with the Chinese loess stratigraphy and the marine isotope record[8,9,21]. Layers that predominantly consist of breakdown breccia (6, 8/9) or loess (4) may be correlated with colder, drier periods, whereas intervening layers (3, 5, 7 and 10) consisting of waterlain sediments and/or flowstones may correspond with warmer, more humid periods (Fig. 1). Following the loess timescale[22], the well-dated flowstone in layer 5 is probably associated with palaeosol 5 (S5) and MIS 13–15 from ~500–600 kyr ago, a period of prolonged warmth and humidity. Continuing downwards, the layer 6 breccia best corresponds to loess 6 (L6) and MIS 16. Layer 7 and the uppermost part of layers 8/9, containing waterlain sediments interbedded with flowstone, would correspond to S6 and MIS 17. The lower part of layers 8/9 is breakdown breccia, and would correspond to L7. The waterlain sediments in layer 10 and below, including the Brunhes/Matuyama boundary between layers 13 and 14, would then correspond to S7 and MIS 19 (refs 4 and 22).

The mammalian fauna support this interpretation. They indicate mixed steppe and forest environments with a trend towards increasing grasslands over time. Layer 5 contains unambiguous warm-climate fauna[8,9]. Other layers show a preponderance of cold-climate fauna associated with the breccia and more warm-climate fauna in the intervening sediments[8,9]. Limited oxygen isotope data also support climatic shifts. Teeth from *Equus sanmeniensis* in layers 8/9 and 4 (glacial) have δ^{18}O values 3–4‰ lower than those in layers 10–11 (interglacial)[23], consistent with enhanced monsoon strength that occurred during the colder, drier intervals.

©2009 Macmillan Publishers Limited. All rights reserved

The above climatic correlations indicate that layers 7–10, including the lower cultural level and the first hominin appearance at Locality 1, lie within the range of S6–S7 and MIS 17–19 from ~0.68 to 0.78 Myr ago. The lower cultural level in the breccia of layers 8/9 would thus correspond to L7 and MIS 18. Pending further confirmation, the assignment of these layers to a cooler and drier episode may imply hominin presence at the site through glacial–interglacial cycles. However, L7 and MIS 18 correspond to a relatively mild glacial period, as indicated by both marine $\delta^{18}O$ and soil development within L7 (Fig. 1). Such mild glacial conditions may have been necessary for early *H. erectus pekinensis* to persist in northern China. Together with previous U-series dating of flowstone in layers 1/2, the hominin presence at the site is constrained to a total range of 0.40–0.78 Myr ago.

A reliable chronology is critical for resolving debate over the mode of Middle Pleistocene human evolution in East Asia[24–26]. Previously, the chronology of Chinese sites has been largely based on the U-series and electron spin resonance dating of fossil materials, which are known to be vulnerable to post-burial U migration[14,27]. $^{230}Th/^{234}U$ dating of speleothem calcite[3,28,29] has repeatedly shown that the previous timescale for Middle–Late Pleistocene hominin sites in China may have been underestimated as a whole. The results of this paper show that such a tendency persists beyond the range of mass spectrometric U-series dating. It is foreseeable that $^{26}Al/^{10}Be$ burial dating will be applied to other hominin sites in China and elsewhere, contributing substantially to a robust chronological framework and thereby to a better understanding of human evolution.

METHODS SUMMARY

For quartzose sand samples, several kilograms of sediment were collected. Silt and clay were removed by a water rinse, and carbonates were dissolved in HCl. The remaining quartz-rich sand was sieved to >0.2 mm and leached several times in hot 5% HF/HNO_3 overnight (>12 h) with agitation. Following magnetic and gravimetric separation, the resulting quartz consisted of two populations: a darker-coloured quartz with a high native aluminium concentration, and a lighter-coloured quartz with a lower aluminium concentration. The darker-coloured grains were removed by handpicking. For quartzite artefacts, the samples were thoroughly cleaned in 1% HF/HNO_3 and then crushed to a grain size <0.5 mm. The quartz was further purified by repeated overnight leaching in 1% HF/HNO_3 in an ultrasonic tank.

Purified quartz was dissolved in 5:1 HF/HNO_3, and spiked with ~0.3 mg ^9Be prepared from beryl. An aliquot was taken for aluminium determination by inductively-coupled plasma optical emission spectrometry using the method of standard additions. After evaporation and fuming of HF in H_2SO_4, aluminium and beryllium were separated on ion-exchange columns in 0.4 M oxalic acid, precipitated as hydroxides, and transformed to oxides in a furnace at 1,100 °C. BeO was mixed with niobium and Al_2O_3 with silver for $^{10}Be/^9Be$ and $^{26}Al/^{27}Al$ measurement by accelerator mass spectrometer at PRIME Lab, Purdue University.

Burial ages were calculated following ref. 16. For samples in this paper, post-burial production of cosmogenic nuclides by muons is safely ignored. Production rates are estimated for latitude 39° N and elevation 120 m (ref. 30), and adjusted for a revised ^{10}Be half-life[20].

Received 9 September; accepted 17 December 2008.

1. Wu, R. K., *et al. Multi-disciplinary Study of the Peking Man Site at Zhoukoudian* [in Chinese] (Science Press, 1985).
2. Zhao, S. S. *et al.* in *Multi-disciplinary Study of the Peking Man Site at Zhoukoudian* [in Chinese] (eds, Wu, R. K. *et al.*) 239–240 (Science Press, 1985).
3. Shen, G. J. *et al.* High-precision U-series dating of Locality 1 at Zhoukoudian, China. *J. Hum. Evol.* **41,** 679–688 (2001).
4. Qian, F., Zhang, J. X. & Yin, W. D. in *Multi-disciplinary Study of the Peking Man Site at Zhoukoudian* [in Chinese] (eds, Wu, R. K. *et al.*) 251–255 (Science Press, 1985).
5. Yuan, S. X., Chen, T. M., Gao, S. J. & Hu, Y. Q. U-series dating of fossil bones from Zhoukoudian sites. *Acta Anthrop. Sin.* **10,** 189–193 (1991).
6. Guo, S. L. *et al.* Fission track dating of Layer 4 of the Peking Man site. *Acta Anthrop. Sin.* **10,** 73–77 (1991).
7. Grün, R. *et al.* ESR analysis of teeth from the palaeoanthropological site of Zhoukoudian, China. *J. Hum. Evol.* **32,** 83–91 (1997).
8. Zhou, C. L., Liu, Z. C., Wang, Y. J. & Huang, Q. H. Climatic cycles investigated by sediment analysis in Peking Man's Cave, Zhoukoudian, China. *J. Archaeol. Sci.* **27,** 101–109 (2000).
9. Liu, Z. C. Sequence of sediments at Locality 1 in Zhoukoudian and correlation with loess stratigraphy in northern China and with the chronology of deep-sea cores. *Quat. Res.* **23,** 139–153 (1985).
10. Yang, Z. G. *et al.* in *Multi-disciplinary Study of the Peking Man Site at Zhoukoudian* [in Chinese] (eds, Wu, R. K. *et al.*) 1–85 (Science Press, 1985).
11. Goldberg, P., Weiner, S., Bar-Yosef, O., Xu, Q. Q. & Liu, J. Y. Site formation processes at Zhoukoudian, China. *J. Hum. Evol.* **41,** 483–530 (2001).
12. Dong, W. *et al.* A comparative analysis on the mammalian faunas associated with *Homo erectus* in China. *Acta Anthrop. Sin.* **19** (suppl.), 246–256 (2000).
13. Aigner, J. S. The age of Zhoukoudian Locality 1: the newly proposed ^{18}O correspondences. *Anthropos* **23,** 157–173 (1986).
14. Ludwig, K. R. & Renne, P. R. Geochronology on the Paleoanthropological time scale. *Evol. Anthropol.* **9,** 101–110 (2000).
15. Richards, D. A. & Dorale, J. A. Uranium-series chronology and environmental applications of speleothems. *Rev. Mineral. Geochem.* **52,** 407–460 (2003).
16. Granger, D. E. & Muzikar, P. F. Dating sediment burial with *in situ*-produced cosmogenic nuclides: theory, techniques, and limitations. *Earth Planet. Sci. Lett.* **188,** 269–281 (2001).
17. Granger, D. E. in *In Situ-Produced Cosmogenic Nuclides and Quantification of Geological Processes* (eds Siame, L., Bourlès, D. L. & Brown, E. T.) 1–16 (Geol. Soc. Am. Spec. Pap. 415, 2006).
18. Partridge, T. C., Granger, D. E., Caffee, M. C. & Clarke, R. Lower Pliocene hominid remains from Sterkfontein. *Science* **300,** 607–612 (2003).
19. Carbonell, E. *et al.* The first hominin of Europe. *Nature* **452,** 465–469 (2008).
20. Nishiizumi, K. *et al.* Absolute calibration of ^{10}Be AMS standards. *Nucl. Instr. Meth. B* **258,** 403–413 (2007).
21. Shackleton, N. J., Berger, A. L. & Peltier, W. R. An alternative astronomical calibration of the lower Pleistocene timescale based on ODP site 677. *Trans. R. Soc. Edinb. Earth Sci.* **81,** 251–261 (1990).
22. Zhou, L. P. & Shackleton, N. J. Misleading positions of geomagnetic reversal boundaries in Eurasian loess and implications for correlation between continental and marine sedimentary sequences. *Earth Planet. Sci. Lett.* **168,** 117–130 (1999).
23. Gaboardi, M., Deng, T. & Wang, Y. Middle Pleistocene climate and habitat change at Zhoukoudian, China, from the carbon and oxygen isotopic record from herbivore tooth enamel. *Quat. Res.* **63,** 329–338 (2005).
24. Rightmire, G. P. Human evolution in the Middle Pleistocene: The role of *Homo heidelbergensis. Evol. Anthropol.* **6,** 218–227 (1998).
25. Wu, X. Z. The evolution of humankind in China. *Acta Anthrop. Sin.* **9,** 312–321 (1990).
26. Clark, J. D. *et al.* African *Homo erectus*: old radiometric ages and young Oldowan assemblages in the Middle Awash Valley, Ethiopia. *Science* **264,** 1907–1910 (1994).
27. Pike, A. W. G. & Pettitt, P. B. U-series dating and human evolution. *Rev. Mineral. Geochem.* **52,** 607–630 (2003).
28. Zhao, J. X., Hu, K., Collerson, K. D. & Xu, H. K. Thermal ionization mass spectrometry U-series dating of a hominid site near Nanjing, China. *Geology* **29,** 27–30 (2001).
29. Shen, G. J., Gao, X., Zhao, J. X. & Collerson, K. D. U-series dating of Locality 15 at Zhoukoudian, China: implications for hominid evolution. *Quat. Res.* **62,** 208–213 (2004).
30. Stone, J. O. Air pressure and cosmogenic isotope production. *J. Geophys. Res.* **105,** 23753–23759 (2000).

Acknowledgements This project was jointly supported by the National Natural Science Foundation of China and Wenner-Gren Foundation. Accelerator mass spectrometer ^{26}Al and ^{10}Be measurements were supported by an NSF facility grant to PRIME Lab. We thank Zhoukoudian Site Museum for assistance in sample collection, L. Zhou, T.-L. Ku and C. Norton for discussions, and S. Zhang and the first author's graduate students Q. Shao, H. Wang and L. Zhang for assistance in sample collection and pre-treatment.

Author Contributions G.S. and D.E.G. contributed equally to conceiving the project, organizing fieldwork, interpreting data and preparing the paper. X.G. provided access to the site and quartzite artefacts. Chemistry was performed by B.G. and D.E.G.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to G.S. (gjshen@njnu.edu.cn) or D.E.G. (dgranger@purdue.edu).

©2009 Macmillan Publishers Limited. All rights reserved

LETTERS

Distinct sensory representations of wind and near-field sound in the *Drosophila* brain

Suzuko Yorozu[1,2], Allan Wong[1,2], Brian J. Fischer[1], Heiko Dankert[1,3], Maurice J. Kernan[4], Azusa Kamikouchi[5,6], Kei Ito[5] & David J. Anderson[1,2]

Behavioural responses to wind are thought to have a critical role in controlling the dispersal and population genetics of wild *Drosophila* species[1,2], as well as their navigation in flight[3], but their underlying neurobiological basis is unknown. We show that *Drosophila melanogaster*, like wild-caught *Drosophila* strains[4], exhibits robust wind-induced suppression of locomotion in response to air currents delivered at speeds normally encountered in nature[1,2]. Here we identify wind-sensitive neurons in Johnston's organ, an antennal mechanosensory structure previously implicated in near-field sound detection (reviewed in refs 5 and 6). Using enhancer trap lines targeted to different subsets of Johnston's organ neurons[7], and a genetically encoded calcium indicator[8], we show that wind and near-field sound (courtship song) activate distinct populations of Johnston's organ neurons, which project to different regions of the antennal and mechanosensory motor centre in the central brain. Selective genetic ablation of wind-sensitive Johnston's organ neurons in the antenna abolishes wind-induced suppression of locomotion behaviour, without impairing hearing. Moreover, different neuronal subsets within the wind-sensitive population respond to different directions of arista deflection caused by air flow and project to different regions of the antennal and mechanosensory motor centre, providing a rudimentary map of wind direction in the brain. Importantly, sound- and wind-sensitive Johnston's organ neurons exhibit different intrinsic response properties: the former are phasically activated by small, bi-directional, displacements of the aristae, whereas the latter are tonically activated by unidirectional, static deflections of larger magnitude. These different intrinsic properties are well suited to the detection of oscillatory pulses of near-field sound and laminar air flow, respectively. These data identify wind-sensitive neurons in Johnston's organ, a structure that has been primarily associated with hearing, and reveal how the brain can distinguish different types of air particle movements using a common sensory organ.

We observed that *Drosophila* exhibit a rapid and reversible arrest of walking activity under gentle air currents (0.7–1.6 m s^{-1}; Fig. 1a, b and Supplementary Movie 1). This behaviour is also exhibited by wild-caught *Drosophila* species at wind speeds (1.7–2.8 m s^{-1}) within the range measured in their natural habitats[1,2,4] (J. S. Johnston, personal communication; Supplementary Information footnote 1). This behaviour, called wind-induced suppression of locomotion (WISL), was observed in the presence or absence of mechanical startle applied to enhance locomotor activity, before the introduction of air flow (Figs 1b and 3d). Importantly, suppression of locomotion was not observed in response to near-field sound stimuli such as courtship song (280 Hz pulse song, 75–100 dB[9]; Supplementary Fig. 1a).

Recent antennal-gluing experiments have implicated the antenna, and by extension Johnston's organ (JO), in wind sensation in *Drosophila*[3,10]. Surgical removal of the third antennal segment (a3),

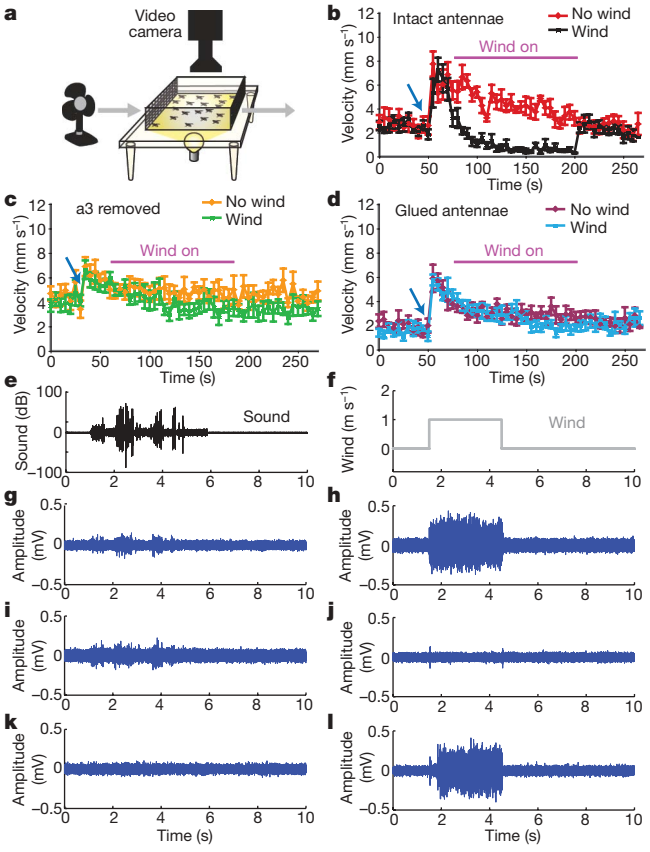


Figure 1 | Behavioural and electrophysiological analyses of wind responses in *Drosophila*. **a**, Schematic illustrating the WISL assay (see Supplementary Methods). **b**, WISL behaviour in Canton-S (CS) flies (Supplementary Movie 1). Data represent mean (± s.e.m.) velocities ($n = 6$). Blue arrow indicates a brief mechanical startle. The 'No wind' versus 'Wind' curves are significantly different ($P = 0.0001$, Kruskal–Wallis analysis of variance (ANOVA)). **c**, **d**, Elimination of WISL by removal of a3 (**c**) or gluing a3 to a2 (**d**). The 'No wind' versus 'Wind' curves are not significantly different ($n = 6$). Data represent mean (± s.e.m). **e–l**, Extracellular recordings of JO neuron responses (blue traces) to sound (**e**) or wind (**f**). **g**, **h**, Response to both sound (**g**) and wind (**h**). **i**, **j**, Response to sound (**i**) but not wind (**j**). **k**, **l**, Response to wind (**l**) but not sound (**k**).

[1]Division of Biology 216-76, [2]Howard Hughes Medical Institute, [3]Division of Engineering and Applied Sciences 136-93, California Institute of Technology, Pasadena, California 91125, USA. [4]Department of Neurobiology and Behavior, SUNY Stony Brook, Stony Brook, New York 11794-5239, USA. [5]Institute of Molecular and Cellular Biosciences, University of Tokyo, Yayoi, Bunkyo-ku, Tokyo 113-0032, Japan. [6]Sensory System Laboratory, Institute of Zoology, University of Cologne, 50923 Cologne, Germany.

©2009 Macmillan Publishers Limited. All rights reserved

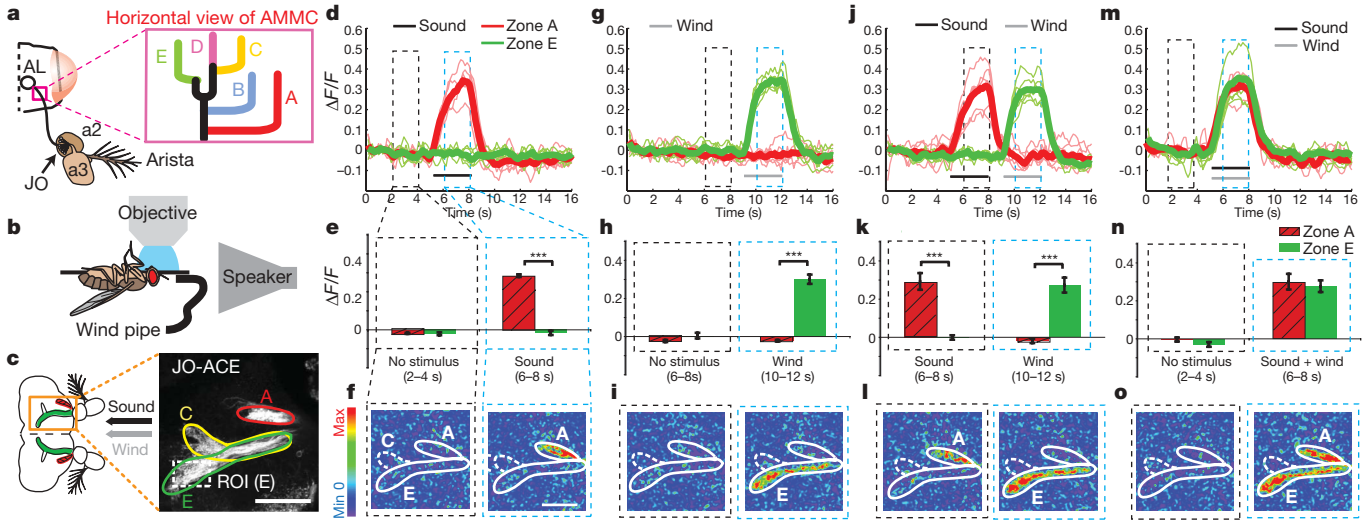


Figure 2 | Calcium imaging reveals distinct populations of wind- and sound-responsive JO neurons. **a–c**, Schematics illustrating location of JO relative to a2 and a3, and five JO neuron axonal terminal zones[7] in the AMMC (**a**), and imaging set-up (**b**). **c**, Zones A, C and E are visualized using a *UAS-mCD8-GFP* reporter. ROI, region of interest for $\Delta F/F$ measurements in zone E. **d–o**, Zones A (red traces, hatched bars) and E (green traces, bars) are activated by sound and wind, respectively, whether presented singly (**d–i**), sequentially (**j–l**) or simultaneously (**m–o**) (see Supplementary Movies 2a–e). Thick traces (**d**, **g**, **j**, **m**) represent the average of the individual (thin) traces ($n = 6$). **e**, **h**, **k**, **n**, Bar graphs indicate the mean (± s.e.m.) integrated $\Delta F/F$ in the time bins (dashed rectangles in **d**, **g**, **j**, **m**; see Methods). $^{***}P < 0.001$ (repeated measure ANOVA and Bonferroni's planned comparisons). **f**, **i**, **l**, **o**, ΔF images of GCaMP activation in zones A and E. Scale bars, 50 μm.

or gluing of a3 to the second antennal segment (a2), both of which cause a functional impairment of JO[11], eliminated WISL (Fig. 1c, d). Genetic ablation of mechanosensory chordotonal neurons using *nanchung-GAL4* (ref. 12) and *UAS-hid* (*head involution defective*, also known as *wrinkled*) a *Drosophila* cell death gene[13], also eliminated WISL (Supplementary Fig. 1b–d). Taken together, these results support the idea that JO is required for WISL, a conclusion confirmed by genetic ablation of specific JO subpopulations (see Fig. 3, later).

To investigate how wind and sound are discriminated by the brain, we first performed extracellular recordings from the antennal nerve[14]. In some electrode placements, spike trains were evoked by both wind (0.3–0.9 m s^{-1}) and courtship song (pulse song; Fig. 1e–h). The short duration of the wind-evoked action potentials (<1 ms) is consistent with neuronal, rather than muscle, action potentials[15]. In other cases, responses were evoked by sound but not wind (Fig. 1i, j; a few spikes were detected at the onset and offset of the wind stimulus), or by wind but not sound (Fig. 1k, l). These results indicate that different axons within the antennal nerve might respond differentially to wind versus sound.

To determine whether distinct subsets of JO neurons are activated by wind versus near-field sound, we performed functional imaging experiments using a genetically encoded calcium sensor (GCaMP-1.3; ref. 8), controlled by different Gal4 enhancer trap lines expressed in JO[7]. These lines identify five major groups of JO axonal projections in the antennal and mechanosensory motor centre (AMMC), called zones A, B, C, D and E (Fig. 2a, inset). Each Gal4 driver labels a subset of zones, but mosaic analysis has revealed that individual JO neurons innervate only one zone[7]. Because it is difficult to distinguish the cell bodies of these five groups of neurons in JO itself, we imaged activity in JO axon terminals in the AMMC, where the five zones are easily discriminated. To do this, we mounted live *Drosophila* in an inverted orientation under a two-photon microscope, while air flow and/or near-field sound were delivered from tubing and a speaker, respectively (Fig. 2b).

Using an enhancer trap line (JO-AB) that selectively labels neurons in zones A and B[7], we observed strong GCaMP activation by courtship song (pulse song; 400 Hz, 90 dB[16]), but not by wind (0.9 m s^{-1}) (Supplementary Fig. 2a–e). Conversely, using a different line (JO-CE) that selectively labels zones C and E[7], we observed responses to air flow, but not to courtship song (Supplementary Fig. 2f–j). To compare directly responses to wind and sound in the same preparation, we used a third line, which labels neurons in zones A, C and E[7] (Fig. 2c). These experiments confirmed that zone A was activated by sound but not by air flow, whereas zone E was activated by air flow but not by sound (Fig. 2d–i and Supplementary Movie 2a, b). The same selective responses were observed when the two stimuli were presented sequentially or simultaneously (Fig. 2j–o and Supplementary Movie 2c, d). Together, these data indicated that JO contains distinct populations of sound- and wind-responsive neurons that project to different regions of the AMMC[7] (Supplementary Information footnote 2).

To determine whether the wind-sensitive JO neurons are also required for WISL behaviour, we genetically ablated these neurons using a toxin, ricin A chain[17]. Because the *JO-CE-GAL4* driver is expressed not only in JO neurons but also in the central brain (Fig. 3a), we used an intersectional strategy to restrict ablation to the antenna using an *eyeless-flippase* (*eyFLP*) tissue-specific recombination system. The specificity of this manipulation was confirmed using an *eyFLP*-dependent mCD8–GFP reporter[18] (Fig. 3b).

Following ablation of JO-C and -E neurons, WISL behaviour was eliminated (Fig. 3g), whereas basal locomotor activity (before wind exposure) and phototaxis behaviour were unaffected (Fig. 3g, i and Supplementary Fig. 3a). Importantly, female flies lacking JO-CE neurons had normal hearing, as evidenced by their unperturbed receptivity to courtship by wild-type males, a behaviour that depends on the females' ability to hear male courtship song. In contrast, females lacking *nanchung*, a gene required for hearing[12], or whose aristae were glued to their head bilaterally[11] (Bi-Gl) exhibited a greatly increased latency to copulation (Fig. 3h, *nan/nan*; Bi-Gl). These data indicate that JO-CE neurons are necessary for WISL behaviour, but dispensable for a hearing-dependent behaviour.

We next investigated the functional significance of the two wind-sensitive JO subpopulations (C and E). Axons innervating zones C and E terminate in lateral versus medial domains of the AMMC, respectively (Fig. 4a–c). When air flow was applied to the front of the head (0°), or at 45°, there was strong activation in zone E and little activation in zone C. Conversely, air flow applied from the rear (180°) activated zone C, and slightly inhibited zone E (Fig. 4d–f and Supplementary Movie 3a–c). Air flow applied to the side of the head

©2009 Macmillan Publishers Limited. All rights reserved

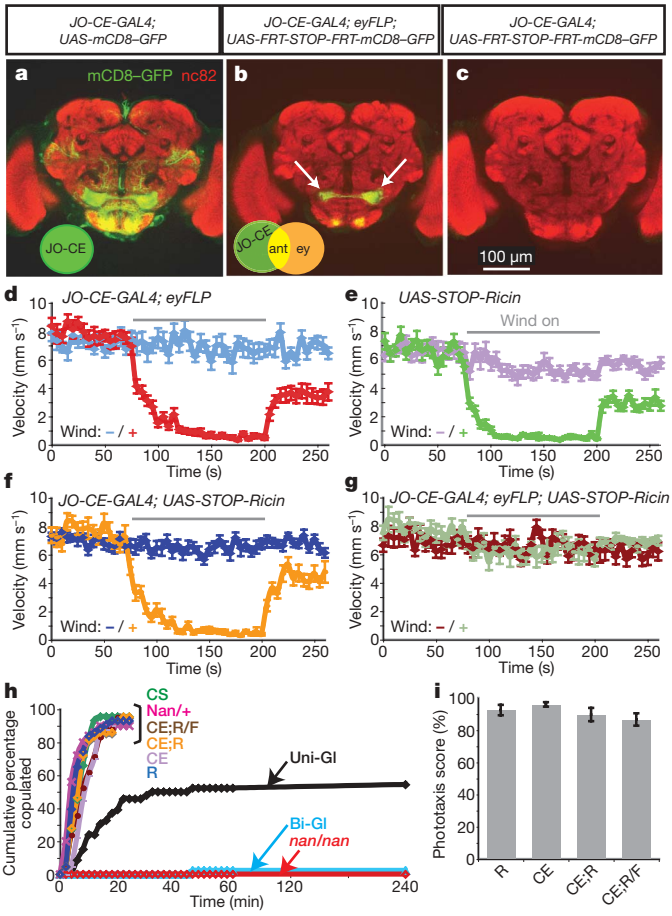


Figure 3 | Ablation of wind-sensitive (C and E) neurons abolishes WISL behaviour. a–c, GFP expression patterns of the indicated genotypes, double-stained with antibodies to GFP (green) and nc82 (red). **a**, Original *JO-CE-GAL4* pattern. **b**, Antennal-restricted *JO-CE-GAL4* pattern using *eyFLP*. Arrows indicate AMMC; underlying structures are gustatory neuron projections to the sub-oesophageal ganglia[7]. **c**, Control for **b** lacking *eyFLP*. **d–g**, WISL analysis. Curves represent mean velocity ± s.e.m. ($n = 15$). **d–f**, Genetic controls show robust WISL. The 'wind−' versus 'wind+' curves are significantly different ($P < 0.0001$). **g**, Ablation of C and E neurons abolishes WISL. **h**, Cumulative percentage of manipulated females copulating with CS males. >90% of all pairs in bracketed conditions exhibited successful copulation within 20 min ($n = 40$–50 pairs). Genotypes: CE;R/F, *JO-CE-GAL4; UAS<STOP<Ricin/eyFLP*; CE;R, *JO-CE-GAL4; UAS<STOP<Ricin*; CE, *JO-CE-GAL4; R, UAS<STOP<Ricin*. Unilateral ('Uni-Gl') or bilateral ('Bi-Gl') gluing of the aristae, or the *nanchung* mutation (*nan/nan*)[12] impaired copulation ($P < 0.0001$). Ablation of JO-CE neurons ('CE;R/F') did not impair either copulation (**h**) or phototaxis (**i**). Phototaxis scores represent mean ± s.e.m. ($n = 7$). *P* values are shown by Kruskal–Wallis ANOVA.

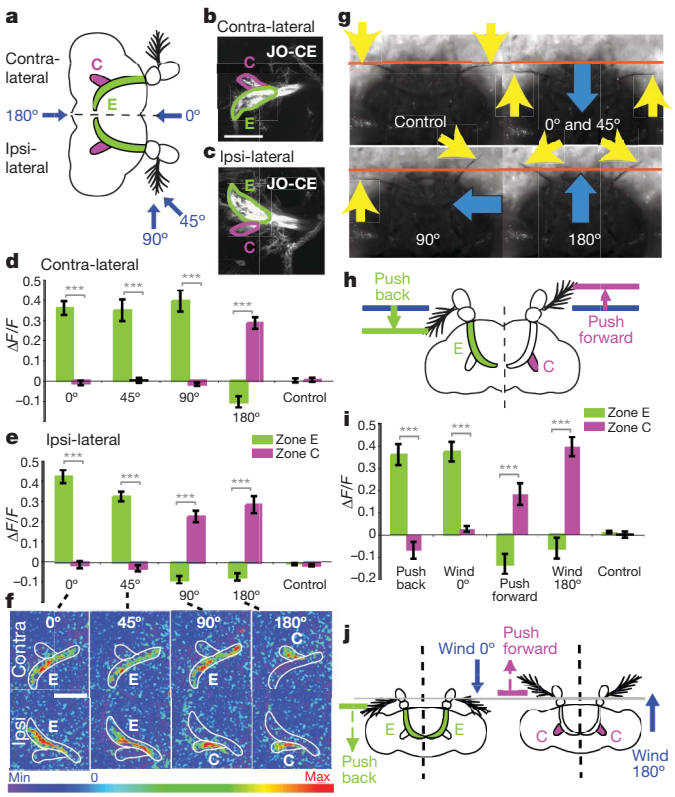


Figure 4 | Wind-direction-sensitivity of zones C versus E. a–c, Schematic (**a**) and mCD8–GFP expression (**b**, **c**) illustrating zones C and E in contra- and ipsi-lateral hemi-brains. Blue arrows indicate wind direction. **d–f**, Average (± s.e.m.) $\Delta F/F$ signals integrated over the stimulus period for zones C and E in the contra- (**d**) and ipsi- (**e**) lateral hemi-brains, and corresponding ΔF images (**f**; see Supplementary Movies 3a–e). **g**, Still frames from video recordings of aristae movements during wind stimulation (Supplementary Movies 4a–e). Yellow arrows indicate aristae position, orange lines denote the rest position ('Control') and blue arrows indicate wind direction. **h**, Schematic illustrating predicted responses of zones C and E to directional, probe-driven aristae displacements. **i**, Responses of the C and E neurons to wind and directional arista displacements. Error bars are s.e.m. **j**, Summary illustrating differential sensitivity of zones C and E to direction of arista displacement. Scale bars, 50 μm. ***$P < 0.0001$ (repeated measures ANOVA and Bonferroni's planned comparisons).

(90°) activated zone C ipsilaterally and zone E contralaterally (Fig. 4d–f, 90°; Supplementary Movie 3d). Thus, zone C and E neurons are differentially sensitive to air flow directionality.

High-magnification video analysis (Supplementary Movie 4a–c) revealed that air flow from different directions moves the aristae either anteriorly or posteriorly (Fig. 4g). We hypothesized that the direction of arista deflection determines whether zone C or E neurons are activated. Arista ablation experiments indicated that the activation of wind-sensitive JO neurons, like that of sound-sensitive JO neurons[11,19], is dependent on this structure (Supplementary Fig. 4). To test the hypothesis directly, we moved the aristae in different directions using a probe controlled by a DC motor (Fig. 4h). Displacing the arista posteriorly with a probe activated the E zone almost as strongly as wind delivered from the front, and weakly inhibited the C zone (Fig. 4i, 'Push back'), whereas displacing it anteriorly activated the C zone and inhibited the E zone (Fig. 4i, 'Push forward'). These data demonstrate that zones C and E are sensitive to different directions of arista deflection (Fig. 4j). This model can explain the asymmetric activation of zones C and E in ipsi- and contral-lateral hemi-brains during wind stimulation from 90° (Fig. 4f, g, 90°), because this stimulus produces opposite deflection of the aristae on the ipsi- and contra-lateral sides of the head (Fig. 4g, 90°, Supplementary Movie 4d). An internal comparison of activity between zones C and E, both within and between each hemi-brain, could provide a basis for computing wind direction[3].

We investigated which stimulus features are responsible for the selective activation of sound- versus wind-sensitive neurons in JO. We first asked whether these two classes of mechanoreceptors are sensitive to different stimulus amplitudes, that is, air particle velocities (v_{air}). A pressure gradient microphone positioned at the antenna[9] yielded a v_{air} of 0.011 m s^{-1} for the 400 Hz sound stimulus played at 90 dB, which maximally activated JO-AB neurons (Fig. 5a). However, this sound stimulus did not activate zone E neurons (Supplementary Fig. 2g), even though these neurons are activated by air flow at a v_{air} as low as 0.005 m s^{-1} (Fig. 5b). Thus, the selectivity of JO-CE and -AB neurons for wind versus sound is not simply due to differences in stimulus magnitude.

We next asked whether JO-AB and -CE neurons might have different intrinsic sensitivities to different types of arista movements by moving

©2009 Macmillan Publishers Limited. All rights reserved

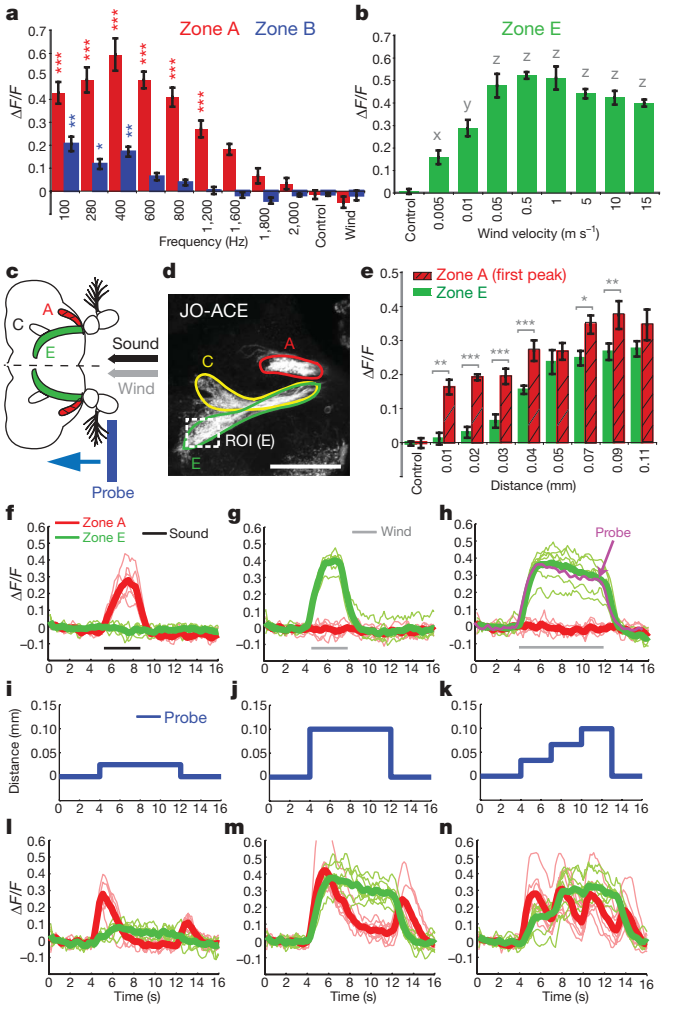


Figure 5 | Wind- and sound-sensitive JO neurons have different intrinsic response properties. **a**, Sensitivity of zones A and B to different sound frequencies (mean ± s.e.m., $n = 6$). ***$P < 0.0001$ (red), *$P < 0.01$ (blue) and **$P < 0.001$ (blue) relative to control. **b**, Sensitivity of zone E to different wind speeds ($n = 5$). Letters (x, y, z) indicate significant differences relative to control (all $P < 0.0001$ except 'x' ($P < 0.001$)). **c–n**, Comparison of sound-, wind- and probe-evoked responses in zones A and E. **c**, **d**, Schematic (**c**) and mCD8–GFP expression (**d**) illustrating zones A, C and E. ROI, region of interest for $\Delta F/F$ measurements in zone E. Scale bar, 50 μm. **e**, Responses of zones A and E to probe-induced arista displacements of different magnitudes. *$P < 0.01$; **$P < 0.001$; ***$P < 0.0001$ (zones A versus E comparisons). All zone A responses, $P < 0.0001$ (relative to control), except for 0.01 mm displacement ($P < 0.001$). All zone E responses ≥0.04 mm displacement, $P < 0.0001$; zone E responses <0.04 mm were not significantly different from control. **f–h**, Sound and wind responses. Thick lines represent average of the individual (thin) traces ($n = 6$). **h**, Superposition of the average responses of zone E to 8 s of wind (green trace) and mechanical probe displacement (magenta trace; see **m**). **i–n**, Responses (**l–n**) of zones A and E to different distances and patterns of probe-induced arista displacement (**i–k**) ($n = 6$). All P values are calculated by repeated measures ANOVA and Bonferroni multiple comparisons.

the aristae in steps of different magnitudes and patterns using a probe controlled by a DC motor (Fig. 5c, d). Sound-sensitive neurons in zone A (Fig. 5f, red traces) were activated by displacements as small as 0.01 mm (Fig. 5e, red hatched bars), whereas wind-sensitive neurons in zone E (Fig. 5g, green traces) were only weakly activated at displacements below 0.04 mm (Fig. 5e, green bars). Thus, zone A neurons have a lower displacement threshold than zone E neurons (see also Fig. 5i, l).

Strikingly, we observed that zone E neurons remained active for as long as the aristae were displaced, whereas zone A neurons were only transiently activated at the onset and offset of probe displacement (Fig. 5j, m). This suggested that zone E neurons might adapt slowly, and therefore respond tonically, whereas zone A neurons might adapt rapidly, and therefore respond phasically. To confirm this, we moved the aristae in three successive steps of 0.033 mm each (total displacement of 0.099 mm; Fig. 5k). Zone A neurons exhibited transient (phasic) responses after each displacement (Fig. 5n, red traces), whereas zone E neurons were tonically activated for the duration of each displacement, and were maximally activated after the second step (Fig. 5n, green traces). These data indicate that JO-AB and JO-CE neurons respond phasically and tonically to arista displacement, with low versus high activation thresholds, respectively (see Supplementary Information footnote 3). Furthermore, zone A neurons were activated by bi-directional movements, whereas zone E neurons were activated only unidirectionally (Fig. 5j, m). These different intrinsic response properties are well matched to the oscillatory arista movements caused by pulses of near-field sound versus uni-directional arista deflections caused by wind. The ability of flies to discriminate wind versus sound using a common sensory organ is thus explained by different populations of JO neurons with different intrinsic response properties, which project to distinct areas of the AMMC.

The identification of different subpopulations of JO neurons with tonic versus phasic response properties illustrates a general and conserved feature of mechanosensation. In mammalian skin, slowly adapting, tonically activated Merkel cells[20] and rapidly adapting, phasically activated Meissner's corpuscles[21] are used for different types of light-touch sensation. In *Drosophila*, these two properties have been adapted to detect different types of bulk air particle movements by different subsets of JO neurons. In the accompanying paper[22], the authors demonstrate, using complementary imaging methods, that zone AB neurons are activated by sound and required for hearing. They also show that zone CE neurons are required for the behavioural response to gravity (negative gravitaxis), a force that could also produce static deflections of the arista, albeit of a smaller magnitude than those produced by wind (Supplementary Information footnote 4).

The data presented here indicate that JO is not simply a hearing organ[6] but also mediates wind detection, in a direction-sensitive manner. Wind-activated neurons in JO are, moreover, required for an innate behavioural response to wind. The function of WISL in nature is not clear. Field studies have suggested that wind is a major environmental factor affecting the dispersal of wild *Drosophila* populations[1,2,4]. WISL may have evolved to control population dispersal, and thereby maintain genetic homogeneity[1,2]. Alternatively, WISL may represent a defence mechanism that serves to protect individual flies from injury, or to prevent dispersal from food resources. Identification of the sensory neurons that mediate WISL opens the way to a systematic analysis of the genes and neural circuitry that underlie this robust, innate behavioural response to wind.

METHODS SUMMARY

Behavioural assay. Twenty flies were used for each WISL trial. A standard WISL trial lasts for 270 s. During the first 55 s, the flies' baseline locomotor activity was recorded. Where indicated, at $t = 55$ s, a brief mechanical stimulus was applied to transiently increase the flies' locomotor activity. Air flow exposure was then initiated at $t = 80$ s, and terminated at $t = 200$ s.

Electrophysiology. Sample preparation and electrophysiological recordings from JO axons were performed as described[14].

Calcium-response imaging. Flies were anaesthetized in a plastic vial on ice for 15−20 s, and then gently inserted into a hole of a thin plastic rectangular plate. After stabilizing the fly with a small drop of wax (55 °C), the proboscis and the area surrounding the proboscis were surgically removed, in a saline bath, to expose the ventral side of the brain. The preparation was then mounted on a microscope in an inverted orientation for calcium-response imaging. The antennae were kept intact and dry throughout the exposure to different stimuli (sound, wind and mechanical probe displacement).

Detailed descriptions of fly stocks, the WISL behavioural apparatus and assay, courtship and phototaxis assays, antennal manipulations, electrophysiology, calcium-response imaging and sample preparation, sound and wind stimuli and statistical methods are provided in the Supplementary Methods.

©2009 Macmillan Publishers Limited. All rights reserved

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 25 November 2008; accepted 29 January 2009.

1. Johnston, J. & Templeton, A. in *Ecological Genetics and Evolution* (eds Barker, J. S. F. & Starmer, W. T.) 241–256 (Academic, 1982).
2. Johnston, J. & Heed, W. Dispersal of desert-adapted *Drosophila*: the saguaro-breeding *D. nigrospiracula*. *Am. Nat.* **110,** 629–651 (1976).
3. Budick, S. A., Reiser, M. B. & Dickinson, M. H. The role of visual and mechanosensory cues in structuring forward flight in *Drosophila melanogaster*. *J. Exp. Biol.* **210,** 4092–4103 (2007).
4. Richardson, R. & Johnston, J. Behavioral components of dispersal in *Drosophila mimica*. *Oecologia* **20,** 287–299 (1975).
5. Caldwell, J. C. & Eberl, D. F. Towards a molecular understanding of *Drosophila* hearing. *J. Neurobiol.* **53,** 172–189 (2002).
6. Kernan, M. J. Mechanotransduction and auditory transduction in *Drosophila*. *Pflügers Arch.* **454,** 703–720 (2007).
7. Kamikouchi, A., Shimada, T. & Ito, K. Comprehensive classification of the auditory sensory projections in the brain of the fruit fly *Drosophila melanogaster*. *J. Comp. Neurol.* **499,** 317–356 (2006).
8. Nakai, J., Ohkura, M. & Imoto, K. A high signal-to-noise Ca^{2+} probe composed of a single green fluorescent protein. *Nature Biotechnol.* **19,** 137–141 (2001).
9. Göpfert, M. C. & Robert, D. The mechanical basis of *Drosophila* audition. *J. Exp. Biol.* **205,** 1199–1208 (2002).
10. Mamiya, A. *et al.* Neural representations of airflow in *Drosophila* mushroom body. *PLoS ONE* **3,** e4063 (2008).
11. Manning, A. Antennae and sexual receptivity in *Drosophila melanogaster* females. *Science* **158,** 136–137 (1967).
12. Kim, J. *et al.* A TRPV family ion channel required for hearing in *Drosophila*. *Nature* **424,** 81–84 (2003).
13. Wang, S. L. *et al.* The *Drosophila* caspase inhibitor DIAP1 is essential for cell survival and is negatively regulated by HID. *Cell* **98,** 453–463 (1999).
14. Eberl, D. F., Hardy, R. W. & Kernan, M. J. Genetically similar transduction mechanisms for touch and hearing in *Drosophila*. *J. Neurosci.* **20,** 5981–5988 (2000).
15. Tanouye, M. A. & Wyman, R. J. Motor outputs of giant nerve fiber in *Drosophila*. *J. Neurophysiol.* **44,** 405–421 (1980).
16. Bennet-Clark, H. C. Acoustics of insect song. *Nature* **234,** 255–259 (1971).
17. Moffat, K. G. *et al.* Inducible cell ablation in *Drosophila* by cold-sensitive ricin A chain. *Development* **114,** 681–687 (1992).
18. Wong, A. M., Wang, J. W. & Axel, R. Spatial representation of the glomerular map in the *Drosophila* protocerebrum. *Cell* **109,** 229–241 (2002).
19. Ewing, A. W. The antenna of *Drosophila* as a 'love song' receptor. *Physiol. Entomol.* **3,** 33–36 (1978).
20. Ikeda, I. *et al.* Selective phototoxic destruction of rat Merkel cells abolishes responses of slowly adapting type I mechanoreceptor units. *J. Physiol. (Lond.)* **479,** 247–256 (1994).
21. Hoffmann, J. N., Montag, A. G. & Dominy, N. J. Meissner corpuscles and somatosensory acuity: the prehensile appendages of primates and elephants. *Anat. Rec. A Discov. Mol. Cell. Evol. Biol.* **281,** 1138–1147 (2004).
22. Kamikouchi, A. *et al.* The neural basis of *Drosophila* gravity-sensing and hearing. *Nature* doi:1010.1038/nature07810 (this issue).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank U. Heberlein and F. Wolf for hosting a sabbatical that led to the discovery of WISL; J. S. Johnson for helpful discussions; L. Zelnik, M. Reiser and P. Perona for creating locomotor tracking software; D. Eberl and J. Hall for *D. melanogaster* courtship song recordings; G. Maimon for making fly holders for imaging experiments; M. Roy for building behavioral chambers for WISL and female receptivity assays; H. Inagaki for *JO-CE-GAL4;eyFLP* flies; B. Hay for *UAS-hid* flies; D. Berdnik for *UAS-FRT-STOP-FRT-Ricin* flies; M. Dickinson for anemometers and discussions; J. L. Anderson for advice on fluid mechanics; M. Göpfert for providing a pressure gradient microphone; M. Konishi for advice and use of laboratory facilities; and G. Mosconi for laboratory management. D.J.A. is an Investigator of the Howard Hughes Medical Institute. This work was supported in part by NSF grant EF-0623527.

Author Contributions S.Y. and D.J.A. designed experiments, S.Y. carried out all experiments reported in this paper and D.J.A. and S.Y. wrote the manuscript. A.W. wrote Matlab programs for $\Delta F/F$ measurements and mechanical probe actuation, B.J.F. assisted with computational filtering of song stimuli, H.D. assisted with computational and statistical analysis of data, M.J.K. provided facilities and support for electrophysiological experiments, and A.K. and K.I. provided Gal4 lines.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to D.J.A. (mancusog@caltech.edu) or S.Y. (yorozu@caltech.edu).

©2009 **Macmillan Publishers Limited. All rights reserved**

METHODS

Fly stocks. Flies carrying *JO-ACE*, *JO-CE* and *JO-AB* were described previously[7]. *UAS-GCaMP*[23,24] and *UAS-mCD8-GFP* flies were obtained from Y. Wang and R. Axel, *UAS-FRT-STOP-FRT-Ricin* flies[25] were obtained from D. Berdnik, *JO-CE-GAL4; eyFLP* flies were obtained from H. Inagaki, Canton-S flies from J. Dubnau, and *UAS-hid* flies from B. Hay. Flies were maintained on corn meal and molasses food at 25 °C on a (12/12) light–dark cycle.

WISL behavioural apparatus. The WISL assay was performed in a $6 \times 6 \times 1.5$ cm transparent acrylic plastic box (WISL chamber), which has air flow inputs and outputs (1 cm diameter) on two of the four vertical sides of the box. The input tubing carries air flow from a tank containing breathable air, connected to a flow regulator. The output tubing allows the air flow to escape from the box, and is connected to a flow meter that measures the speed of the air flow. The WISL chamber was mounted on a transparent plastic table and was trans-illuminated with a fluorescent light from underneath. A video camera (Sony, DCR-HC40 NTSC) was set up above the WISL chamber to record the flies' locomotor activity.

WISL assay protocol. Twenty flies per trial were sorted 36–48 h before testing, using nitrogen gas or cold anaesthesia. On the testing day, 20 flies were aspirated into the WISL chamber and allowed to acclimate for 7–8 min just before initiation of the trial. A standard WISL trial lasted for 270 s. Flies were given a brief mechanical stimulation (5 manual strikes on the table that the WISL chamber was mounted on) at 55 s, and air flow exposure began at 80 s and ended at 120 s. Locomotor activity was recorded at ten frames per second and average velocity was computed using custom software written in Matlab (MathWorks Inc.).

Courtship (female receptivity) assay. Naive Canton-S males and virgin females of the genotype of interest were collected immediately after eclosion, using nitrogen or CO_2 gas anaesthesia. Naive males were individually housed whereas virgin females were group-housed for six days until the test day. Single naive Canton-S male and a single virgin female of the genotype of interest were placed in a mating chamber ($1 \times 1 \times 0.4$ cm square chamber), and the time at which a successful copulation occurred was recorded for each mating pair. Successful copulation typically lasts 15–25 min.

Phototaxis assay. Forty flies per trial were sorted 48 h before testing, using nitrogen or CO_2 gas anaesthesia. On the test day, 40 flies were tapped into the elevator of a T-maze and allowed to rest for one minute in a dark. Next, the elevator was lowered to the choice point where flies were given one minute to make a choice between a dark tube and a tube illuminated with a 40 W fluorescent light, positioned approximately 20 cm away. The phototaxis response was analysed by calculating the performance index (PI) using the following formula: $PI = [(2 \times COR) - 1] \times 100$, where COR = (number of flies that chose the illuminated tube/total number of flies). PI = 0 indicates an equal distribution of flies between the dark and illuminated tubes. PI = 100% indicates that all flies chose the illuminated tube.

Antenna manipulations. To test the role of JO in wind detection, a3 segments were surgically removed using a pair of forceps, 48 h before the WISL testing. For the antennal gluing experiment, a small drop of ultraviolet-activated glue was placed at the junction between the a2 and a3 segments bilaterally, and cured with an ultraviolet lamp for 3–5 s, 48 h before the testing. For the mechanical probe antennal displacement experiment, a sharpened tungsten needle was used to move the aristae in different directions and different patterns. The probe was mounted on a DC motor/controller (LTA-HS and SMC100CC, Newport), which was controlled by custom Matlab software (MathWorks Inc.). To push the aristae backward, the probe was positioned anterior to the aristae; conversely, to push the aristae forward, the probe was positioned posterior to the aristae. In the 'push back' (and 'push forward') conditions, the aristae were pushed backward (or forward) in a single increment of varying distances (0.01, 0.02, 0.025, 0.03, 0.04, 0.05, 0.07, 0.09 or 0.11 mm), held for 8 s in the displaced position and then returned to the original position. In another experiment, the aristae was pushed backwards in three successive steps of 0.033 mm (a total of 0.099 mm), held in place for 2.9 s after each successive step, and then returned to the original position. In all conditions, the probe and aristae movements were verified using a video camera (GE680, Proscillica) that was set up underneath the fly preparation mounted on the microscope stage as described previously.

Electrophysiology. Extracellular recordings from JO axons were recorded at the a1/a2 joint using a tungsten or glass electrode (0.5 MΩ) as described previously[14] in a sound-proof chamber. Pulse-song segments of recorded *D. melanogaster* courtship song (provided by J. Hall[26] and D. Eberl) were used as the sound stimulus and an air flow rate between 0.3 and 0.9 m s^{-1} was used as the wind stimulus.

Calcium-response imaging and sample preparation. Flies were anaesthetized in a plastic vial on ice for ~15–20 s, and were then gently inserted into a hole of a thin plastic rectangular plate. Small drops of wax (55 °C) were applied to prevent the fly from moving out from the hole. After the fly was stabilized in the plastic hole, the preparation was oriented in an upside-down position (see Fig. 2b). The proboscis, ventral part of thorax and abdomen, and legs were protruding from the upper side of the horizontal plane of the plastic, while the rest of the fly head (including the antennae, but excluding the proboscis), thorax and dorsal part of abdomen were protruding from the bottom side of the horizontal plane of the plastic. In a saline bath, the proboscis was cut off and the area surrounding the proboscis was surgically removed to expose the ventral side of the brain. Fat and air sacs were gently removed to give a clear view of the brain. For calcium-response imaging, the water immersion objective lens (×40, NA = 0.8, Olympus) was lowered near the exposed brain, while the underside of the plastic specimen mount, which contained the intact antennae, was kept dry and exposed to wind and sound stimuli.

Sound stimuli used in these experiments were recorded segments (provided by J. Hall[26] and D. Eberl) of the pulse-song portion of *D. melanogaster* courtship song, played at 75–100 dB at the aristae using a loudspeaker (ProMonitor 800 loudspeaker, Definitive Technology) and amplifier (P.A. amplifier, Radioshack) and were measured using a digital sound meter (DSM-325, Mannix). We tested the frequency tuning of zones A and B using narrowband signals derived from the original pulse-song. The original pulse-song was filtered to set the centre of the frequency spectrum at a desired frequency between 100 and 2,000 ± 200 Hz (using the Fourier transformation).

Wind stimuli used in imaging experiments were delivered at speeds between 0.005 and 15 m s^{-1}. Wind speed was controlled by a flow regulator (mass flow meters and controllers, Smart Trak series 100, Sierra Instrument Inc.) and was measured using an anenometer (Testo-435, Testo GmbH & Co.). VClamp software (Pairie Technology) was used to control all aspects of sound and wind stimuli used in the imaging experiments.

All imaging was performed on an Ultima two-photon laser scanning microscope (Prairie Technology). Live images were acquired at 6.1 frames per second using an Olympus ×40 (NA = 0.8) water immersion objective at 128×128 resolution with an imaging wavelength at 925 nm. GCaMP responses were quantified using custom software written in MatLab. The relative change in fluorescence intensity ($\Delta F/F$) was computed by first calculating the average pixel values in the region of interest during the experimental period and applying a three-frame moving average smoothing function. This average fluorescence value, F_{av}, was then converted to $\Delta F/F$ using the formula $\Delta F/F = (F_{av} - F_0)/F_0$, where F_0 is the baseline fluorescence value, measured as the average of frames 2–9. Average $\Delta F/F$ for a specific stimulus period was compared between different JO neuron zones to test for statistical significance by repeated-measure ANOVA.

23. Wang, J. W., Wong, A. M., Flores, J., Vosshall, L. B. & Axel, R. Two-photon calcium imaging reveals an odor-evoked map of activity in the fly brain. *Cell* **112,** 271–282 (2003).
24. Wang, Y. *et al.* Stereotyped odor-evoked activity in the mushroom body of Drosophila revealed by green fluorescent protein-based Ca^{2+} imaging. *J. Neurosci.* **24,** 6507–6514 (2004).
25. Berdnik, D., Chihara, T., Couto, A. & Luo, L. Wiring stability of the adult *Drosophila* olfactory circuit after lesion. *J. Neurosci.* **26,** 3367–3376 (2006).
26. Wheeler, D. A., Fields, W. L. & Hall, J. C. Spectral analysis of *Drosophila* courtship songs: *D. melanogaster*, *D. simulans*, and their interspecific hybrid. *Behav. Genet.* **18,** 675–703 (1988).

©2009 Macmillan Publishers Limited. All rights reserved

LETTERS

Enhancing SIV-specific immunity *in vivo* by PD-1 blockade

Vijayakumar Velu[1,2]*, Kehmia Titanji[1,2]*, Baogong Zhu[3,4], Sajid Husain[1,2], Annette Pladevega[1,2], Lilin Lai[1,2], Thomas H. Vanderford[5], Lakshmi Chennareddi[1,2], Guido Silvestri[5], Gordon J. Freeman[3,4], Rafi Ahmed[1] & Rama Rao Amara[1,2]

Chronic immunodeficiency virus infections are characterized by dysfunctional cellular and humoral antiviral immune responses[1–3]. As such, immune modulatory therapies that enhance and/or restore the function of virus-specific immunity may protect from disease progression. Here we investigate the safety and immune restoration potential of blockade of the co-inhibitory receptor programmed death 1 (PD-1)[4,5] during chronic simian immunodeficiency virus (SIV) infection in macaques. We demonstrate that PD-1 blockade using an antibody to PD-1 is well tolerated and results in rapid expansion of virus-specific CD8 T cells with improved functional quality. This enhanced T-cell immunity was seen in the blood and also in the gut, a major reservoir of SIV infection. PD-1 blockade also resulted in proliferation of memory B cells and increases in SIV envelope-specific antibody. These improved immune responses were associated with significant reductions in plasma viral load and also prolonged the survival of SIV-infected macaques. Blockade was effective during the early (week 10) as well as late (~week 90) phases of chronic infection even under conditions of severe lymphopenia. These results demonstrate enhancement of both cellular and humoral immune responses during a pathogenic immunodeficiency virus infection by blocking a single inhibitory pathway and identify a novel therapeutic approach for control of human immunodeficiency virus infections.

Virus-specific T cells show varying degrees of functional impairment during chronic infections[6,7]. Although these T cells retain some antiviral functions, they are less polyfunctional compared with antiviral T cells seen in acute infections. This defect in T-cell function greatly contributes to the inability of the host to eliminate the persisting pathogen. The exhaustion of virus-specific T cells was first shown during persistent lymphocytic choriomeningitis virus (LCMV) infection of mice[8,9] and was quickly extended to other model systems including human immunodeficiency virus (HIV), hepatitis B virus (HBV) and hepatitis C virus (HCV) infections in humans[1,2,10]. The co-inhibitory receptor PD-1 has been shown to be highly expressed by exhausted LCMV-specific CD8 T cells and *in vivo* blockade of this pathway restores function in these cells[11]. PD-1 is also upregulated on HIV-1-specific[12–14] and SIV-specific[15,16] CD8 T cells and *in vitro* blockade of PD-1 enhances cytokine production and proliferative capacity of these cells. However, the importance of this PD-1 inhibitory pathway in regulating T-cell function during immunodeficiency virus infection *in vivo* is not known. Here, we use a SIV/macaque model to evaluate the effects of *in vivo* blockade of PD-1 on the safety and restoration of virus-specific cellular and humoral immunity during chronic immunodeficiency virus infections.

PD-1 blockade was performed using an antibody specific to human PD-1 that blocks the interaction between macaque PD-1 and its ligands (PDLs) *in vitro*[15]. Blockade was performed during the early (10 weeks) as well as late (~90 weeks) phases of chronic SIV infection. Nine macaques (five during the early phase and four during the late phase) received the anti-PD-1 antibody and five macaques (three during the early phase and two during the late phase) received an isotype control antibody (Synagis, anti-respiratory syncytial virus (RSV)-specific)[17].

PD-1 blockade during chronic SIV infection resulted in a rapid expansion of SIV-specific CD8 T cells in the blood of all macaques (Fig. 1a, b). We were able to study the CD8 T-cell responses to two immunodominant epitopes, Gag CM9 (ref. 18) and Tat SL8/TL8 (ref. 19), using major histocompatibility complex (MHC) I tetrameric complexes in seven of the anti-PD-1-antibody-treated and three of the control-antibody-treated macaques that expressed the Mamu A*01 histocompatibility molecule. Consistent with previous data[15], most (>98%) of the Gag-CM9 tetramer-specific CD8 T cells expressed PD-1 before blockade (data not shown). After PD-1 blockade, the Gag-CM9 tetramer-specific CD8 T cells expanded rapidly and peaked by 7–21 days. At the peak response, these levels were about 2.5–11-fold higher than their respective levels on day 0 ($P = 0.007$) and remained elevated until 28–45 days (Fig. 1b). Similar results were observed with blockade during the early as well as late phases of chronic SIV infection. A 3–4-fold increase in the frequency of Gag-specific interferon (IFN)-γ-positive CD8 T cells was also observed by day 14 after blockade in the two Mamu A*01-negative animals (RTd11 and RDb11), demonstrating that PD-1 blockade can enhance the frequency of virus-specific CD8 T cells that are restricted by non-Mamu A*01 alleles (data not shown). As expected, expansion of SIV-specific CD8 T cells was not observed in the control-antibody-treated macaques (Fig. 1).

PD-1 blockade was also associated with a significant increase in the frequency of virus-specific CD8 T cells that were undergoing active cell division *in vivo* with improved functional quality (Fig. 1b). Consistent with the rapid expansion of SIV-specific CD8 T cells, the frequency of Gag-CM9 tetramer-specific CD8 cells that co-expressed Ki67 (marker for proliferating cells) also increased as early as by day 7 after blockade ($P = 0.01$). Similarly, we observed an increase in the frequencies of Gag-CM9 tetramer-specific CD8 T cells co-expressing perforin and granzyme B (cytolytic potential; $P = 0.001$ and $P = 0.03$, respectively), CD28 (co-stimulation potential; $P = 0.001$), CD127 (proliferative potential; $P = 0.0003$)[20] and CCR7 (lymph-node homing potential; $P = 0.001$)[21]. We also observed a transient 1.5–2-fold increase in the

[1]Emory Vaccine Center, Emory University School of Medicine, Atlanta, Georgia 30322, USA. [2]Division of Microbiology and Immunology, Yerkes National Primate Research Center, Emory University, Atlanta, Georgia 30329, USA. [3]Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, Massachusetts 02115, USA. [4]Department of Medicine, Harvard Medical School, Boston, Massachusetts 02115, USA. [5]University of Pennsylvania School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA. *These authors contributed equally to this work.

©2009 Macmillan Publishers Limited. All rights reserved

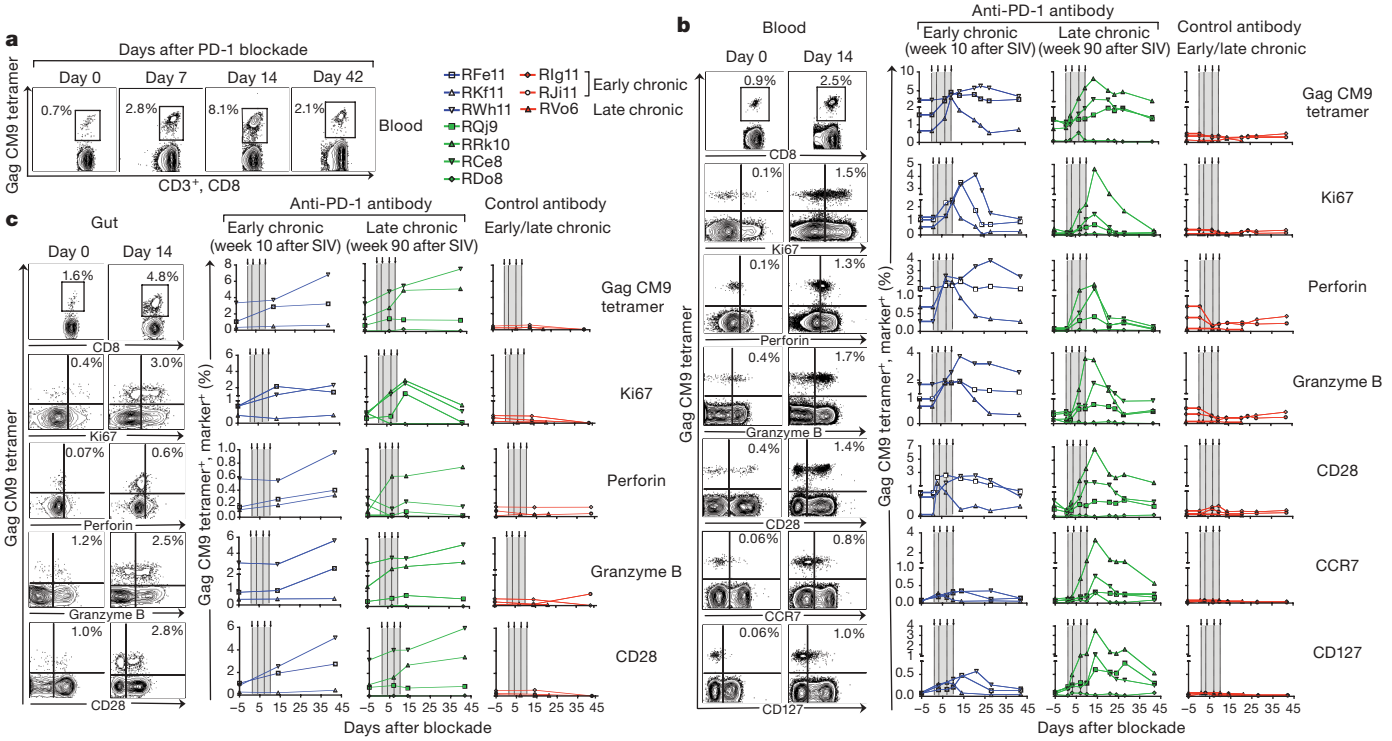


Figure 1 | *In vivo* PD-1 blockade during chronic SIV infection increases the Gag-CM9-specific CD8 T cells with improved functional quality in both blood and gut. **a**, Representative FACS plots for macaque RRk10. **b**, **c**, The magnitude and phenotype of Gag-CM9-tetramer-positive CD8 T cells in blood (**b**) and gut (colorectal mucosal tissue) (**c**). Representative FACS plots are shown on the left and summary for all Mamu A*01-positive animals is shown on the right. Numbers on the FACS plots represent the frequency of tetramer-positive cells as a per cent of total CD8 T cells. Arrows and vertical lines indicate anti-PD-1 antibody or control antibody treatment.

frequency of tetramer-negative and Ki67-positive CD8 T cells after blockade (data not shown). This could be due to expansion of CD8 T cells specific to other epitopes in Gag as well as other proteins of SIV, and other chronic viral infections in these animals. No significant enhancement was observed for these markers in the three control-antibody-treated macaques.

Notably, no expansion was observed for Tat-TL8-specific CD8 T cells after blockade (Supplementary Fig. 1a). This could be due to viral escape from recognition by Tat-TL8-specific CD8 T cells, as PD-1 blockade is known to result in expansion of T cells only when they simultaneously receive signals through T-cell receptor. To test this possibility, we sequenced the viral genomes present in the plasma just before the initiation of blockade from all three Mamu A*01-positive macaques that were infected with SIV251 and received the blocking antibody during the early phase of infection. Indeed, we found mutations in the viral genome corresponding to the Tat TL8 epitope region (Supplementary Fig. 1b). All these mutations either have been shown or predicted to reduce the binding of Tat SL8/TL8 peptide to Mamu A*01 MHC molecule and result in escape from recognition by the Tat-SL8/TL8-specific CD8 T cells[18,19]. These results suggest that *in vivo* blockade of PD-1 may not result in expansion of T cells that are specific to escape mutants of viral epitopes.

PD-1 blockade also resulted in expansion of Gag-CM9-specific CD8 T cells at the colorectal mucosal tissue (gut), a preferential site of SIV/HIV replication[22] (Fig. 1c). Expansion was not observed for two of the seven macaques, although expansion was evident for one of them in blood. In contrast to blood, the expansion in gut peaked much later by day 42 and ranged from 2- to 3-fold compared with their respective day 0 levels ($P = 0.003$). Similar to blood, the Gag-CM9 tetramer-specific cells that co-expressed Ki67 ($P = 0.01$), perforin ($P = 0.03$), granzyme B ($P = 0.01$) and CD28 ($P = 0.01$) also increased in the gut after blockade.

More importantly, PD-1 blockade also enhanced the functional quality of anti-viral CD8 T cells and resulted in the generation of polyfunctional cells capable of co-producing the cytokines IFN-γ, tumour-necrosis factor (TNF)-α and interleukin (IL)-2 (Fig. 2). On the day of initiation of PD-1 blockade during the late chronic phase of infection, the frequency of Gag-specific IFN-γ-positive cells was low and they failed to co-express TNF-α and IL-2 (Fig. 2a). However, after the blockade, the frequency of IFN-γ-positive cells increased in all four PD-1 antibody-treated macaques ($P = 0.03$) and they acquired the ability to co-express TNF-α and IL-2. The expansion of IFN-γ-positive cells peaked by 14–21 days and the peak levels were 2–10-fold higher than the respective day 0 levels. On day 21, about 16% of the total Gag-specific cells co-expressed all three cytokines, and about 30% co-expressed IFN-γ and TNF-α (Fig. 2b). This is in contrast to $<1\%$ of the total Gag-specific cells co-expressing all three cytokines ($P = 0.01$), and about 14% co-expressing IFN-γ and TNF-α on day 0 ($P = 0.04$). Similar results were also observed after blockade during the early chronic phase of infection (data not shown).

Recent studies have shown that chronic immunodeficiency virus infections are also associated with B-cell dysfunction[3] and very little is known about the role of PD-1 in regulating B-cell function/exhaustion. To understand the role of PD-1 in regulating B-cell function during chronic immunodeficiency virus infections, we characterized the B-cell responses after PD-1 blockade in SIV-infected macaques (Fig. 3). Analysis of PD-1 expression on different B-cell subsets before PD-1 blockade revealed preferential expression of PD-1 by memory B cells ($CD20^+CD27^+CD21^-$) compared to naive B cells ($CD20^+CD27^-CD21^+$; Fig. 3a, $P < 0.001$). *In vivo* blockade of PD-1 resulted in a 2–8-fold increase in the titre of SIV-specific binding antibody by day 28 after blockade ($P < 0.001$; Fig. 3b). To understand this further, we studied the proliferation of memory B cells in SIV-infected macaques that were treated simultaneously with anti-PD-1 antibody and anti-retroviral therapy and observed a significant increase in $Ki67^+$ (proliferating) memory, but not naive, B cells as early as day 3 (Fig. 3c). These results demonstrate that the PD-1–PDL pathway could have a role in regulating B-cell dysfunction during chronic SIV infection.

©2009 Macmillan Publishers Limited. All rights reserved

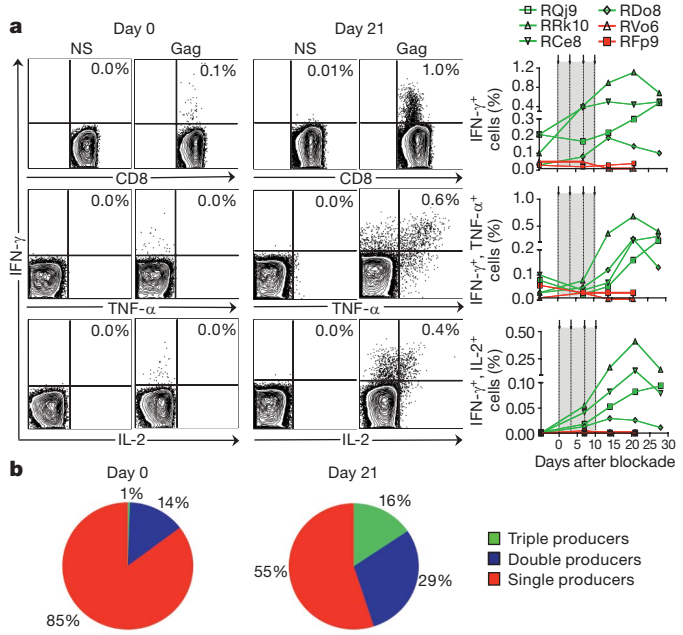


Figure 2 | *In vivo* PD-1 blockade during chronic SIV infection increases the polyfunctional virus-specific CD8 T cells. a, Frequency of Gag-specific cytokine-secreting CD8 T cells as a percentage of total CD8 T cells. Representative FACS plots are shown on the left and summary for the group is shown on the right. Arrows and vertical lines indicate anti-PD-1 antibody or control antibody treatment. Green lines represent anti-PD-1-antibody-treated macaques and red lines represent control-antibody-treated macaques. NS, no stimulation. **b**, Cytokine co-expression subsets expressed as a percentage of total cytokine-positive cells. Mean for the group is shown. Single producers, IFN-γ^+, TNF-α^+ or IL-2$^+$; double producers, IFN-γ^+TNF-α^+, IFN-γ^+IL-2$^+$ or IL-2$^+$TNF-α^+; triple producers, IFN-γ^+IL-2$^+$TNF-α^+.

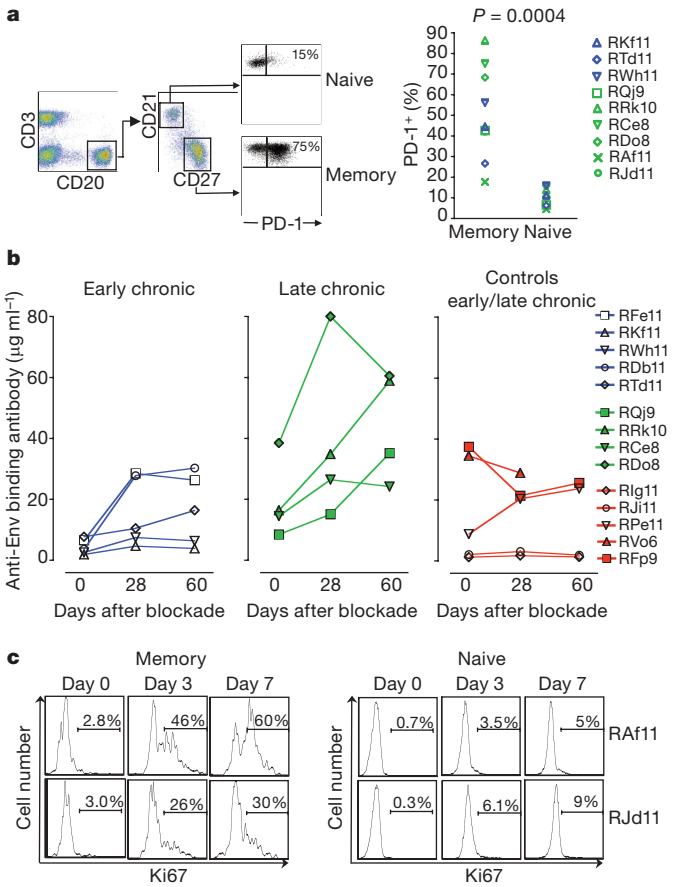


Figure 3 | *In vivo* PD-1 blockade during chronic SIV infection enhances SIV-specific humoral immunity. a, Expression of PD-1 on memory (CD20$^+$CD27$^+$CD21$^-$) and naive (CD20$^+$CD27$^-$CD21$^+$) B cells in blood after SIV infection and before *in vivo* PD-1 blockade. **b**, Titres of anti-SIV Env-binding antibody in serum after blockade. **c**, Ki67 expression (marker for proliferation) on memory and naive B cells after blockade. Numbers on the FACS plots represent Ki67-positive cells as a percentage of respective total naive/memory B cells. Macaques RAf11 and RJd11 were treated simultaneously with anti-PD-1 antibody and anti-retroviral therapy at 22 weeks after SIV infection.

Neutralization assays revealed a twofold increase in titres against the easily neutralizable laboratory-adapted SIV251 and no increase in titres against hard-to-neutralize wild-type SIV251 or SIV239 (data not shown). In two of the nine animals treated with anti-PD-1 antibody, we observed only a minimal (<2-fold) expansion of SIV-specific antibody after blockade. Notably, the frequency of total memory B cells in these two animals was lower (~40% of total B cells) compared with the remaining seven animals (60–90% of total B cells) before blockade (data not shown), indicating that the level of SIV-specific memory B cells before blockade may determine the level of expansion of SIV-specific antibody after blockade.

PD-1 blockade resulted in significant reductions in plasma viraemia ($P = 0.03$) and also prolonged the survival of SIV-infected macaques ($P = 0.001$; Fig. 4). In two of the five macaques treated with anti-PD-1 antibody during the early chronic phase, viral load declined by day 10 and persisted at or below this level until day 90 (Fig. 4a). In one macaque viral load declined transiently and in the remaining two macaques increased transiently and returned to pre-blockade levels. In contrast to the early chronic phase, all four macaques treated with the anti-PD-1 antibody during the late chronic phase showed a transient increase in viraemia by day 7, but rapidly reduced the virus load by day 21 to levels that were below their respective day 0 levels (Fig. 4b). However, the viral RNA levels returned to pre-blockade levels by day 43. As expected, no significant reductions in the plasma viral loads were observed in any of the five macaques treated with the control antibody (Fig. 4c). By 21–28 days after blockade, the viral RNA levels in the anti-PD-1-antibody-treated animals were 2–10-fold lower than their respective day 0 levels ($P = 0.03$; Fig. 4d). By day 150 after the blockade, four of the five macaques in the control group were killed owing to AIDS-related symptoms (for example loss of appetite, diarrhoea, weight loss), whereas all nine animals in the anti-PD-1-antibody-treated group had survived ($P = 0.001$; Fig. 4e).

The observed initial rise in plasma viraemia levels in all of the late-phase-treated and some of the early-phase-treated animals could be due to an increase in the frequency of activated CD4 T cells. To determine this, we measured the percentage of Ki67-positive total CD4 T cells as well as the frequency of SIV Gag-specific IFN-γ-producing CD4 T cells (preferential targets for virus replication[23]) after blockade (Supplementary Fig. 2). These analyses revealed a transient increase in the percentage of Ki67-positive CD4 T cells by day 7–14 after blockade ($P = 0.002$) and this increase was higher in animals treated during the late phase than early phase of infection ($P = 0.015$; Supplementary Fig. 2a). Similarly, an increase in the frequency of Gag-specific CD4 T cells was also observed, but only in animals treated during the late phase of infection (Supplementary Fig. 2b). No significant increases were observed for these activated CD4 T cells in the control-antibody-treated macaques. These results suggest that the activated CD4 T cells could have contributed to the observed initial rise in plasma viraemia levels after blockade.

It is to be noted that, before initiation of PD-1 blockade, the set point viral load in plasma and total CD4 T cells in blood and gut were similar between the anti-PD-1-antibody-treated and control-antibody-treated groups (Supplementary Fig. 3). However, the frequencies of Gag CM9$^+$ cells and Gag CM9$^+$ cells co-expressing perforin, granzyme B or CD28 were not similar between the two treatment groups before *in vivo* blockade (Fig. 1b). This raises the possibility that

©2009 Macmillan Publishers Limited. All rights reserved

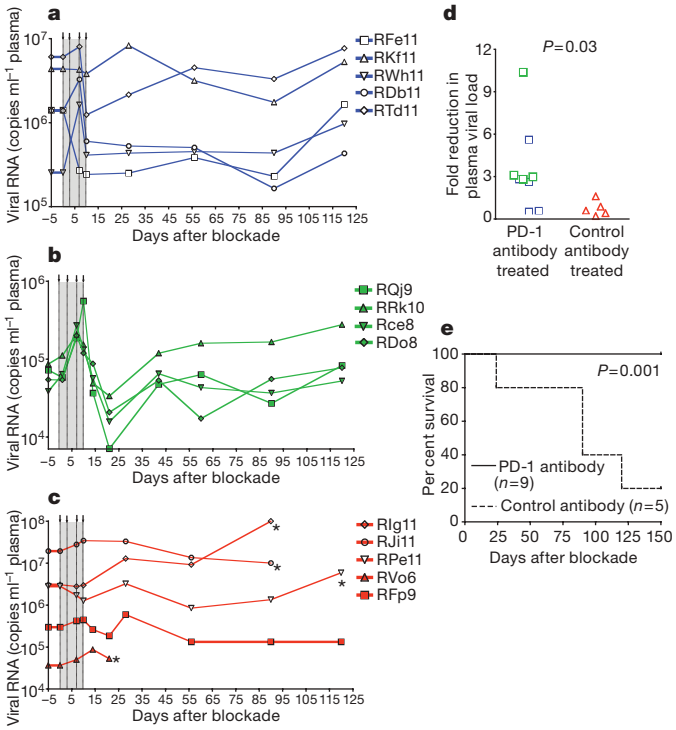


Figure 4 | *In vivo* PD-1 blockade reduces plasma viraemia and prolongs survival of SIV-infected macaques. **a**–**c**, Plasma viral load in macaques treated with anti-PD-1 antibody during the early chronic phase of infection (**a**), macaques treated with anti-PD-1 antibody during the late chronic phase of infection (**b**), and macaques treated with control antibody during the early/late chronic phase of SIV infection (**c**). An asterisk indicates death of animal. **d**, Fold reduction in plasma viral load between day 0 and day 28 (early chronic study) or day 0 and day 21 (late chronic study). The colour scheme is the same as that used in **a**–**c**. **e**, Survival of SIV-infected macaques after PD-1 blockade.

these differences could have contributed to the expansion of Gag $CM9^+$ cells after PD-1 blockade. To study the influence of the frequency of Gag $CM9^+$ cells before blockade on their expansion after blockade, we divided the anti-PD-1-antibody-treated group into two subgroups based on the frequency of Gag $CM9^+$ cells before initiation of blockade such that one group has similar levels and the other group has higher levels of Gag $CM9^+$ cells compared with the control-antibody-treated group. These subgroups were then analysed for expansion of Gag $CM9^+$ cells after blockade. Expansion of Gag $CM9^+$ cells was evident in both subgroups of animals after blockade of PD-1, irrespective of whether they were at low or high levels before blockade (Supplementary Fig. 4). Similar results were also observed with subgroup analyses based on the frequency of Gag $CM9^+$ cells co-expressing molecules associated with better T-cell function such as perforin, granzyme B, CCR7, CD127 or CD28 (Supplementary Fig. 4). However, we observed a trend towards better expansion of Gag $CM9^+CD28^+$ cells in animals with higher levels of Gag $CM9^+CD28^+$ cells before blockade, suggesting that CD28 expression may serve as a biomarker for predicting the outcome of *in vivo* PD-1 blockade.

To evaluate the safety of PD-1 blockade, we performed an extensive analysis of serum proteins, ions, lipids, liver and kidney enzymes, and complete blood count after blockade (Supplementary Tables 1 and 2). These analyses revealed no significant changes for all parameters tested between the anti-PD-1-antibody-treated and control-antibody-treated macaques (data not shown). Similarly, the levels of anti-nuclear antibodies (ANA) in serum (measure of autoimmunity) did not change significantly after treatment with anti-PD-1 antibody (Supplementary Fig. 5). In one macaque, the levels of ANA increased about 3-fold by day 10 after blockade, but returned to day 0 levels by day 56. These results demonstrate that anti-PD-1 antibody treatment during chronic SIV infection results in no observable toxicity. This is consistent with a recent study that demonstrated the safety of PD-1 blockade in patients with advanced haematological malignancies[24].

We studied the pharmacokinetics of the partially humanized anti-PD-1 antibody in serum after *in vivo* blockade. The titre of anti-PD-1 antibody rapidly declined between days 14 and 28 after blockade (Supplementary Fig. 6a) and coincided with macaques generating antibody response against the mouse immunoglobulin variable domains of anti-PD-1 antibody (Supplementary Fig. 6b). These results suggest that the use of completely humanized anti-PD-1 antibody may allow longer periods of treatment that may further enhance the efficacy of *in vivo* blockade.

Our results demonstrate that *in vivo* blockade of PD-1 during chronic SIV infection is safe and results in rapid expansion and restoration of SIV-specific polyfunctional CD8 T cells and enhanced B-cell responses. Expansion was observed with blockade performed during the early as well as late phases of chronic infection even under conditions of high levels of persisting viraemia and AIDS. Expansion was also observed at the colorectal mucosal tissue, a preferential site of SIV/HIV replication[22]. Importantly, PD-1 blockade resulted in a significant reduction of plasma viral load and also prolonged the survival of SIV-infected macaques. These results are highly significant considering the failure of blockade of a related co-inhibitory molecule CTLA-4 to expand virus-specific CD8 T cells and to reduce plasma viral load in SIV-infected macaques[25]. The therapeutic benefits of PD-1 blockade may be improved further by using combination therapy with anti-retrovirals and/or therapeutic vaccination.

METHODS SUMMARY

Study group. Fourteen Indian rhesus macaques (*Macaca mulatta*) infected with SIV were studied. Eight macaques were used for the early chronic phase and were infected intravenously with 200 50% tissue culture infectious dose ($TCID_{50}$) of SIV251. Six macaques were used for the late chronic phase, three were infected with SIV251 intrarectally and three were infected with SIV239 intravenously. All macaques, except RDb11, were negative for Mamu B08 and Mamu B17 alleles. RDb11 was positive for Mamu B17 allele. Macaques were housed at the Yerkes National Primate Research Center and were cared for under guidelines established by the Animal Welfare Act and the NIH 'Guide for the Care and Use of Laboratory Animals' using protocols approved by the Emory University IACUC.

***In vivo* antibody treatment.** Macaques were infused with either partially humanized mouse anti-human PD-1 antibody (clone EH12-1540)[26] or a control antibody (SYNAGIS). The anti-PD-1 antibody has mouse variable heavy chain domain linked to human IgG1 (mutated to reduce FcR and complement binding)[27] and mouse variable light chain domain linked to human κ. The clone EH12 binds to macaque PD-1 and blocks interactions between PD-1 and its ligands *in vitro*[15]. SYNAGIS is a humanized mouse monoclonal antibody (IgG1κ) specific to F protein of respiratory syncytial virus (Medimmune). Antibodies were administered intravenously at 3 mg kg^{-1} of body weight on days 0, 3, 7 and 10.

Immune responses. Peripheral blood mononuclear cells from blood and lymphocytes from rectal pinch biopsies were isolated as described previously[15]. Tetramer staining[28], intracellular cytokine production[28,29] and measurements of anti-SIV Env binding antibody[30] were performed as described previously.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 14 October; accepted 21 November 2008.
Published online 10 December 2008.

1. Letvin, N. L. & Walker, B. D. Immunopathogenesis and immunotherapy in AIDS virus infections. *Nature Med.* **9**, 861–866 (2003).
2. Pantaleo, G. & Koup, R. A. Correlates of immune protection in HIV-1 infection: what we know, what we don't know, what we should know. *Nature Med.* **10**, 806–810 (2004).
3. Moir, S. & Fauci, A. S. Pathogenic mechanisms of B-lymphocyte dysfunction in HIV disease. *J. Allergy Clin. Immunol.* **122**, 12–19 (2008).
4. Ishida, Y., Agata, Y., Shibahara, K. & Honjo, T. Induced expression of PD-1, a novel member of the immunoglobulin gene superfamily, upon programmed cell death. *EMBO J.* **11**, 3887–3895 (1992).
5. Sharpe, A. H., Wherry, E. J., Ahmed, R. & Freeman, G. J. The function of programmed cell death 1 and its ligands in regulating autoimmunity and infection. *Nature Immunol.* **8**, 239–245 (2007).

©2009 Macmillan Publishers Limited. All rights reserved

6. Wherry, E. J. *et al.* Molecular signature of $CD8^+$ T cell exhaustion during chronic viral infection. *Immunity* **27,** 670–684 (2007).
7. Klenerman, P. & Hill, A. T cells and viral persistence: lessons from diverse infections. *Nature Immunol.* **6,** 873–879 (2005).
8. Zajac, A. J. *et al.* Viral immune evasion due to persistence of activated T cells without effector function. *J. Exp. Med.* **188,** 2205–2213 (1998).
9. Gallimore, A. *et al.* Induction and exhaustion of lymphocytic choriomeningitis virus-specific cytotoxic T lymphocytes visualized using soluble tetrameric major histocompatibility complex class I-peptide complexes. *J. Exp. Med.* **187,** 1383–1393 (1998).
10. Rehermann, B. & Nascimbeni, M. Immunology of hepatitis B virus and hepatitis C virus infection. *Nature Rev. Immunol.* **5,** 215–229 (2005).
11. Barber, D. L. *et al.* Restoring function in exhausted CD8 T cells during chronic viral infection. *Nature* **439,** 682–687 (2006).
12. Petrovas, C. *et al.* PD-1 is a regulator of virus-specific $CD8^+$ T cell survival in HIV infection. *J. Exp. Med.* **203,** 2281–2292 (2006).
13. Day, C. L. *et al.* PD-1 expression on HIV-specific T cells is associated with T-cell exhaustion and disease progression. *Nature* **443,** 350–354 (2006).
14. Trautmann, L. *et al.* Upregulation of PD-1 expression on HIV-specific $CD8^+$ T cells leads to reversible immune dysfunction. *Nature Med.* **12,** 1198–1202 (2006).
15. Velu, V. *et al.* Elevated expression levels of inhibitory receptor programmed death 1 on simian immunodeficiency virus-specific CD8 T cells during chronic infection but not after vaccination. *J. Virol.* **81,** 5819–5828 (2007).
16. Petrovas, C. *et al.* SIV-specific $CD8^+$ T-cells express high levels of PD1 and cytokines but have impaired proliferative capacity in acute and chronic SIVmac251 infection. *Blood* **110,** 928–936 (2007).
17. Malley, R. *et al.* Reduction of respiratory syncytial virus (RSV) in tracheal aspirates in intubated infants by use of humanized monoclonal antibody to RSV F protein. *J. Infect. Dis.* **178,** 1555–1561 (1998).
18. Allen, T. M. *et al.* Characterization of the peptide binding motif of a rhesus MHC class I molecule (Mamu-A*01) that binds an immunodominant CTL epitope from simian immunodeficiency virus. *J. Immunol.* **160,** 6062–6071 (1998).
19. Allen, T. M. *et al.* Tat-specific cytotoxic T lymphocytes select for SIV escape variants during resolution of primary viraemia. *Nature* **407,** 386–390 (2000).
20. Kaech, S. M. *et al.* Selective expression of the interleukin 7 receptor identifies effector CD8 T cells that give rise to long-lived memory cells. *Nature Immunol.* **4,** 1191–1198 (2003).
21. Sallusto, F., Lenig, D., Forster, R., Lipp, M. & Lanzavecchia, A. Two subsets of memory T lymphocytes with distinct homing potentials and effector functions. *Nature* **401,** 708–712 (1999).
22. Pierson, T., McArthur, J. & Siliciano, R. F. Reservoirs for HIV-1: mechanisms for viral persistence in the presence of antiviral immune responses and antiretroviral therapy. *Annu. Rev. Immunol.* **18,** 665–708 (2000).
23. Douek, D. C. *et al.* HIV preferentially infects HIV-specific $CD4^+$ T cells. *Nature* **417,** 95–98 (2002).
24. Berger, R. *et al.* Phase I safety and pharmacokinetic study of CT-011, a humanized antibody interacting with PD-1, in patients with advanced hematologic malignancies. *Clin. Cancer Res.* **14,** 3044–3051 (2008).
25. Cecchinato, V. *et al.* Immune activation driven by CTLA-4 blockade augments viral replication at mucosal sites in simian immunodeficiency virus infection. *J. Immunol.* **180,** 5439–5447 (2008).
26. Dorfman, D. M., Brown, J. A., Shahsafaei, A. & Freeman, G. J. Programmed death-1 (PD-1) is a marker of germinal center-associated T cells and angioimmunoblastic T-cell lymphoma. *Am. J. Surg. Pathol.* **30,** 802–810 (2006).
27. Xu, D. *et al. In vitro* characterization of five humanized OKT3 effector function variant antibodies. *Cell. Immunol.* **200,** 16–26 (2000).
28. Amara, R. R. *et al.* Control of a mucosal challenge and prevention of AIDS by a multiprotein DNA/MVA vaccine. *Science* **292,** 69–74 (2001).
29. Kannanganat, S., Ibegbu, C., Chennareddi, L., Robinson, H. L. & Amara, R. R. Multiple-cytokine-producing antiviral CD4 T cells are functionally superior to single-cytokine-producing cells. *J. Virol.* **81,** 8468–8476 (2007).
30. Lai, L. *et al.* GM-CSF DNA: an adjuvant for higher avidity IgG, rectal IgA, and increased protection against the acute phase of a SHIV-89.6P challenge by a DNA/MVA immunodeficiency virus vaccine. *Virology* **369,** 153–167 (2007).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements The authors thank J. D. Altman for provision of Gag CM9 and Tat SL8 tetramers, H. Drake-Perrow for administrative support, and D. Watkins and Wisconsin National Primate Research Center Genotyping Service for Mamu typing of animals. The authors also thank the Yerkes Division of Research Resources for pathology support, Emory CFAR virology core for viral load assays and the NIH AIDS Research and Reference Reagent Program for the provision of peptides. This work was supported by the National Institutes of Health/National Institute of Allergy and Infectious Diseases grants R01 AI057029, R01 AI071852, R01 AI074417 to R.R.A.; the Foundation for the NIH through the Grand Challenges in Global Health initiative P51 RR00165 to R.A., G.J.F. and R.R.A.; Yerkes National Primate Research Center base grant P51 RR00165; Emory CFAR grant P30 AI050409; and R24 RR16038 to David I. Watkins.

Author Contributions V.V. and K.T. contributed to the design of experiments, conducted analyses on T-cell responses, and contributed to manuscript preparation; S.H. performed analyses on T-cell phenotyping. T.H.V. performed analyses on viral escape. L.L. and A.P. performed analyses on humoral responses; L.C. performed the statistical analysis; G.S. supervised the analyses on viral escape and contributed to manuscript preparation; B.Z. and G.J.F. developed and provided the anti-human PD-1 blocking antibody, and contributed to the design of experiments and manuscript preparation. R.A. contributed to the concept, design of experiments and manuscript preparation. R.R.A. supervised the entire project, designed and coordinated the experiments, and contributed to manuscript preparation.

Author Information Sequencing data related to the Tat SL8/TL8 epitope region have been deposited in GenBank (accession numbers FJ268664–FJ268704). Reprints and permissions information is available at www.nature.com/reprints. The authors declare competing financial interests: details accompany the full-text HTML version of the paper at www.nature.com/nature. Correspondence and requests for materials should be addressed to R.R.A. (ramara@emory.edu).

©2009 **Macmillan Publishers Limited. All rights reserved**

METHODS

B-cell responses. A total of 100 μl of blood was surface stained with antibodies to CD3 (clone SP34-2, BD Biosciences), CD20 (2H7, e-Biosciences), CD21 (B-ly4, Becton Dickson) CD27 (M-T2712, Becton Dickson) and PD-1 (clone EH-12), each conjugated to a different fluorochrome. Cells were lysed and fixed with FACS lysing solution, and permeabilized using FACS perm (BD Biosciences) according to the manufacturer's instructions. Cells were then stained for intracellular Ki67 using an anti-Ki67 antibody conjugated to phycoerythrin (clone B56, Becton Dickson). After staining, cells were washed and acquired using LSRII (BD Biosciences), and analysed using FlowJo software.

Titres of anti-PD-1 antibody and monkey antibody response against anti-PD-1 antibody in serum. To measure the levels of anti-PD-1 antibody, plates were coated with goat anti-mouse immunoglobulin (pre-absorbed to human immunoglobulin, Southern Biotech), blocked and incubated with different dilutions of serum to capture the PD-1 antibody. Bound antibody was detected using anti-mouse IgG conjugated to HRP (pre-absorbed to human immunoglobulin, Southern Biotech). Known amounts of PD-1 antibody captured in the same manner were used to generate a standard curve. To measure the levels of monkey antibody response against the anti-PD-1 antibody, plates were coated with anti-PD-1 antibody (5 $\mu g\,ml^{-1}$), blocked and incubated with different dilutions of serum to capture monkey antibody against the anti-PD-1 antibody. Bound antibody was detected using anti-human λ-chain-specific antibody conjugated to HRP (Southern Biotech). This detection antibody does not bind to the PD-1 antibody because only the constant regions of the heavy and light chains were humanized and the constant region of the light chain is κ. The amount of captured monkey immunoglobulin was estimated using a standard curve that consisted of known amounts of purified macaque immunoglobulin that had been captured using anti-macaque immunoglobulin.

Quantification of SIV copy number. SIV copy number was determined using a quantitative real-time PCR as previously described[28]. All specimens were extracted and amplified in duplicates, with the mean result reported.

Amplification and sequencing of the Tat TL8 epitope. A 350-nucleotide fragment including Tat TL8 epitope was amplified by limiting dilution RT–PCR. Viral RNA was extracted using the QIAamp Viral RNA mini kit (Qiagen) from plasma. vRNA was reverse transcribed with the SIVmac239-specific primer Tat-RT3 (5′-TGGGGATAATTTTACACAAGGC-3′) and Superscript III (Invitrogen) using the manufacturer's protocol. The resultant cDNA was diluted and copy number was determined empirically in our nested PCR protocol. Limiting dilution, nested PCR was performed at ~0.2 copies per reaction using the Expand HiFi PCR kit (Roche Applied Sciences) with outer primers Tat-F1 (5′-GATGAATGGGTAGTGGAGGTTCTGG-3′) and Tat-R2 (5′-CCCAAGTATCCCTATTCTTGGTTGCAC-3′), and inner primers Tat-F3 (5′-TGATCCTCGCTTGCTAACTG-3′) and Tat-R3 (5′-AGCAAGATGGCGATAAGCAG-3′). The first round reactions were cycled using the following programme: 94 °C for 1 min, followed by 10 cycles of 94 °C for 30 s, 55 °C for 30 s, and 68 °C for 1 min, followed by 25 more cycles identical to the first ten but for the addition of 5 s to the extension time at every cycle, followed by a final extension at 68 °C for 7 min. The second round reactions were cycled using the following programme: 94 °C for 1 min, followed by 35 cycles of 94 °C for 30 s, 53 °C for 30 s, and 68 °C for 1 min, followed by a final extension at 68 °C for 7 min. After clean-up with ExoSap-IT (USB Corporation), PCR products were sequenced directly using the inner primers on an automated sequencer at The Children's Hospital of Philadelphia's Napcore sequencing facility. Contigs were assembled using Sequencher 4.8 (Gene Codes Corporation). Amplicons containing nucleotides with double chromatogram peaks were excluded.

Statistical analyses. Linear mixed effects models were used to determine differences in blood chemistry and complete blood count values between anti-PD-1-antibody-treated and control-antibody-treated animals. The Bonferroni method was used to adjust P values for multiple tests. A paired t-test was used for comparison of immune responses before and after PD-1 blockade. Log-transformed data were used when the data were not normal, but log-normal. A Wilcoxon rank-sum test was used to compare the fold reductions in viral loads between the groups. A Mantel–Haenszel log rank test was used to compare the survival curves between the groups. Statistical analyses were performed using S-PLUS 8.0. A two-sided $P < 0.05$ was considered statistically significant.

©2009 Macmillan Publishers Limited. All rights reserved

LETTERS

Complete but curtailed T-cell response to very low-affinity antigen

Dietmar Zehn[1], Sarah Y. Lee[1] & Michael J. Bevan[1]

After an infection, T cells that carry the CD8 marker are activated and undergo a characteristic kinetic sequence of rapid expansion, subsequent contraction and formation of memory cells[1–3]. The pool of naive T-cell clones is diverse and contains cells bearing T-cell antigen receptors (TCRs) that differ in their affinity for the same antigen[4,5]. How these differences in affinity affect the function and the response kinetics of individual T-cell clones was previously unknown. Here we show that during the *in vivo* response to microbial infection, even very weak TCR–ligand interactions are sufficient to activate naive T cells, induce rapid initial proliferation and generate effector and memory cells. The strength of the TCR–ligand interaction critically affects when expansion stops, when the cells exit lymphoid organs and when contraction begins; that is, strongly stimulated T cells contract and exit lymphoid organs later than weakly stimulated cells. Our data challenge the prevailing view that strong TCR ligation is a prerequisite for $CD8^+$ T-cell activation. Instead, very weak interactions are sufficient for activation, but strong TCR ligation is required to sustain T-cell expansion. We propose that in response to microbial challenge, T-cell clones with a broad range of avidities for foreign ligands are initially recruited, and that the pool of T cells subsequently matures in affinity owing to the more prolonged expansion of high-affinity T-cell clones.

So far, our knowledge of the response kinetics of T cells that carry the CD8 marker to pathogens has come from analysing the response of monoclonal TCR transgenic T cells to high-affinity TCR ligands[1–3,6–11]. We set up experiments to compare the polyclonal response of endogenous T cells with that of TCR transgenic T cells at early and late time points after an infection. Three thousand OT-1 TCR transgenic $CD8^+$ T cells, specific for an ovalbumin (Ova) peptide plus H-$2K^b$, were transferred into mice that were subsequently infected with recombinant *Listeria monocytogenes* expressing Ova (Lm-N4ova). On day 4 post-infection the mice contained similar numbers of K^b/Ova-specific endogenous $CD8^+$ cells and OT-1 T cells (Fig. 1a, b). In contrast, by day 7 the ratio had shifted by more than 30-fold in favour of the OT-1 cells (Fig. 1a, b). This shift indicated that the OT-1 cells expanded between days 4 and 7 better than the endogenous T cells. The population of endogenous $CD8^+$ T cells on day 4 was far less sensitive to antigen than were the OT-1 cells (Fig. 1c), but by day 7 both populations were equally sensitive to antigen stimulation (Fig. 1d). In addition, we detected changes in the TCR-β chain usage among endogenous K^b/Ova-specific T cells between days 4.5 and 7.5 post-infection (Supplementary Fig. 1). Thus, although many T cells with low functional avidity are initially recruited into the response, higher-avidity T cells expand better than lower-avidity clones, suggesting a more complex dynamic within the endogenous population than a simple equal and synchronous expansion of all the recruited clones.

Low functional avidity can occur when T cells carry a TCR that recognizes a ligand with low affinity[12]. To study how TCR affinity for antigen impacts clonal propagation in a controlled situation, we chose a model in which TCR transgenic OT-1 cells are stimulated by different altered peptide ligands (APLs). We selected five APLs derived from the original OT-1 ligand SIINFEKL (N4). The APLs bind equally well to H-$2K^b$ as N4 but differ in their potency for stimulating OT-1 cells (Supplementary Fig. 2). To ensure that the APLs are presented to OT-1 cells in a physiological context during an infection, we generated recombinant *L. monocytogenes* strains engineered to express chicken Ova protein containing APLs (Lm-APLova) in place of the N4 epitope.

All of the Lm-APLova strains, even Lm-V4ova, containing the V4 peptide, which is about 700-fold less potent than N4 (Supplementary Fig. 2), expand OT-1 cells *in vivo* (Supplementary Fig. 3). Nonetheless, there was a direct correlation between OT-1 numbers

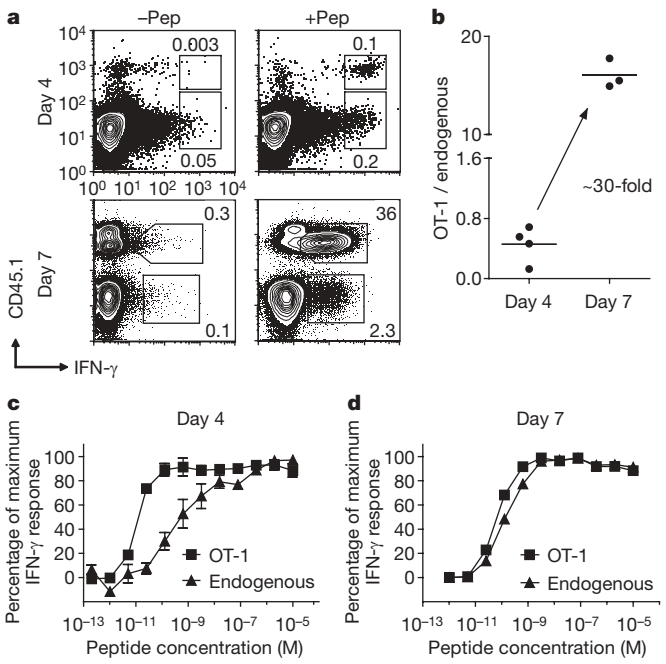


Figure 1 | Unequal propagation of OT-1 and endogenous $CD8^+$ T cells. C57BL/6 mice were grafted with 3×10^3 naive CD45.1 OT-1 cells and infected with *L. monocytogenes* expressing Ova. On days 4 and 7 post-infection, total splenocytes were harvested and briefly re-stimulated with Ova peptide *in vitro*. **a**, Representative $CD8^+$-gated flow plots of splenocytes stimulated with or without 10 μM peptide. Frequencies refer to total $CD8^+$ cells. **b**, Ratios of the numbers of OT-1 and endogenous IFN-γ T cells found in individual mice. **c**, **d**, Peptide dose–response curves normalized to the level of maximum numbers of interferon-γ-producing OT-1 and endogenous K^b/Ova-specific T cells ($n = 4$, day 4; $n = 3$, day 7); bars show s.e.m.

[1]Department of Immunology, Howard Hughes Medical Institute, University of Washington, Box 357370, Seattle, Washington 98195, USA.

©2009 Macmillan Publishers Limited. All rights reserved

and the affinity of the priming ligand on day 6 post-infection (Fig. 2a). To study the initial kinetics of the OT-1 response to Lm-APLova, we grafted mice with higher numbers of carboxy-fluorescein diacetate succinimidyl ester (CFSE)-labelled OT-1 cells and infected them with *Listeria* expressing either the native N4 peptide, or the low-affinity APL, Q4, or the very low-affinity APL, V4. In all cases, at 78 h post-infection, most divided OT-1 cells are in their sixth and seventh division (Fig. 2b). This indicates that the affinity of TCR interaction does not dictate the rate of initial cell division. Lm-N4ova and Lm-Q4ova drive similar accumulation of OT-1 T cells in the spleen 4 days post-infection (Fig. 2c). However, although the N4-stimulated OT-1 cells keep expanding and plateau at day 7 post-infection, Q4-stimulated OT-1 reach their maximum number on day 5.5 and start to contract earlier than N4-stimulated OT-1 (Fig. 2c). Cells stimulated by weaker ligands showed reduced 5-bromodeoxyuridine (BrdU) incorporation at day 5.5 post-infection than N4-primed OT-1 cells, implying that they go through fewer cell divisions rather than dividing and dying (Supplementary Fig. 4). Measuring the frequency of OT-1 cells after infection with each of the different Lm-APLova strains revealed that the weaker the ligand, the earlier the cells reach their maximum level of expansion and begin to contract (Fig. 2d). Thus, although low-affinity ligands initiate the same rate of rapid proliferation as high-affinity ligands, the final magnitude of expansion and the timing of the onset of contraction are determined by the strength of the TCR–ligand interaction. The efficient recruitment of low-affinity $CD8^+$ T cells and their shorter expansion period explain the data presented in Fig. 1.

T cells stimulated by low-affinity ligands, including V4, were phenotypically similar to cells stimulated by high-affinity ligands in expressing CD11b, CD27, CD28, CD44, CD122 and PD-1 on days 4 and 7 post-infection (data not shown). OT-1 cells stimulated by low- and high-affinity ligands expressed granzyme B on day 4 post-infection, lysed antigen-positive target cells *in vivo* and produced interferon-γ (IFN-γ) and tumour-necrosis factor-α (TNF-α) in response to peptide stimulation on day 7 post-infection (Supplementary Fig. 5). The only major phenotypic differences we observed between OT-1 cells stimulated by low- and high-affinity ligands were the surface levels of CD25 and CCR7, which were lower on weakly stimulated OT-1 cells on day 4 post-infection (Fig. 2e). These data reveal that even very weak TCR ligation induces cells with typical effector phenotypes.

After infection with *Listeria*, specific $CD8^+$ T cells meet antigen in the splenic periarteriolar lymphocyte sheaths (PALS) and undergo explosive proliferation before any progeny exit the PALS and enter the red pulp and the blood stream[13,14]. Supplementary Fig. 6 shows that the progeny of a single OT-1 cell stimulated by Lm-N4ova remain in the same PALS until day 4. Surprisingly, OT-1 cells stimulated by weak ligands such as Q4 and V4 can be detected in the blood and in the red pulp at 4 days post-infection, before N4-stimulated OT-1 cells appear in these compartments (Fig. 3). Similarly, in the polyclonal endogenous response, effector T cells detectable in the blood at day 4.5 post-infection are of much lower functional avidity than the cells found at day 7.5 (Supplementary Fig. 7). The early appearance of weakly stimulated cells in the blood and the red pulp correlates with the earlier downregulation of CCR7 on Q4-stimulated OT-1 (Fig. 2e). These data show that ligand affinity dictates the kinetics of T-cell migration and that, remarkably, the earliest wave of effectors released into the blood stream is dominated by T cells bearing low-affinity TCR.

All of the APLs also induce an endogenous $CD8^+$ T-cell response (data not shown). In the case of APLs that interact weakly with the OT-1 TCR, some endogenous $CD8^+$ T-cell clones respond better to the APLs than the OT-1 cells (data not shown). Could competition with endogenous T cells bearing higher-affinity TCR explain the early migration out of the spleen and the shortened expansion time of weakly stimulated OT-1 cells? Against this notion, we found that the response to Lm-Q4ova or Lm-V4ova was not enhanced in TCR-$\alpha^{-/-}$ P14 TCR transgenic mice, which completely lack an endogenous $CD8^+$ T-cell response to Q4 or V4 (Supplementary Fig. 8).

We then asked whether $CD8^+$ T cells stimulated by low-affinity ligands converted to memory cells. We followed the response of OT-1 cells stimulated by the different ligands for 138 days. In all cases, we detected a population of memory OT-1 cells that remained stable in number beyond day 21 (Fig. 4a, b). When challenged with recombinant vesicular stomatitis virus expressing Ova (VSV-N4ova), the

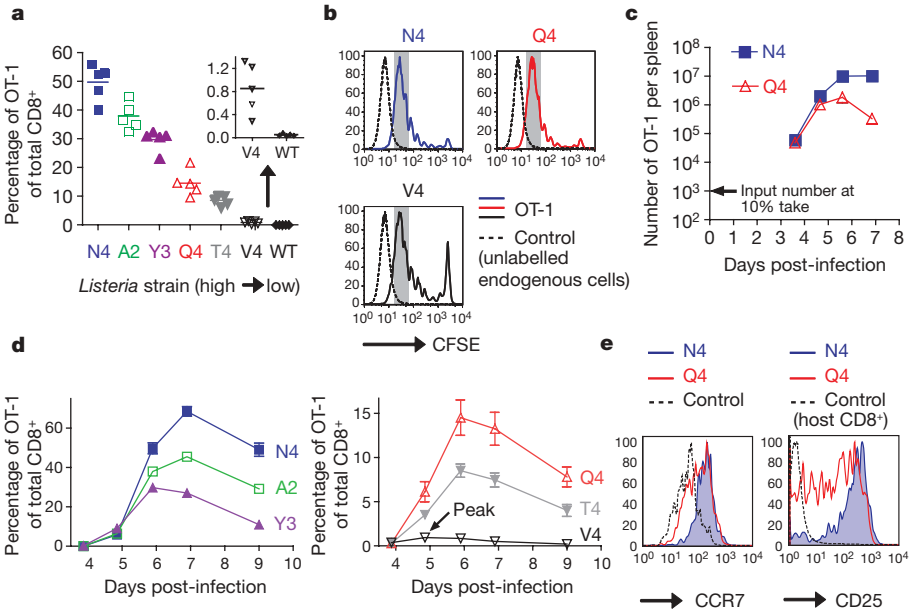


Figure 2 | The strength of TCR ligation dictates the timing of T-cell contraction. Mice were grafted with 10^4 unlabelled (in **a**, **c**–**e**) or 2×10^5 CFSE-labelled (**b**) naive OT-1 cells and infected with wild-type (WT) or recombinant *L. monocytogenes* expressing Ova protein containing native SIINFEKL (N4) or the indicated altered peptide ligands (APLs), listed in order of decreasing potency (Supplementary Fig. 2). **a**, Frequency of OT-1 among $CD8^+$ cells in the blood at day 6 post-infection; inset shows Lm-V4 and wild-type *Listeria* on a magnified scale. **b**, CFSE dilution profiles of splenic OT-1 cells at day 3.25 post-infection. **c**, Numbers of splenic OT-1 cells at the indicated time points ($n = 4$ per group and time point). **d**, Frequency of OT-1 among $CD8^+$ blood cells at the indicated time points ($n = 5$); bars show s.e.m. **e**, Surface CCR7 and CD25 levels on splenic OT-1 at day 4 post-infection.

©2009 Macmillan Publishers Limited. All rights reserved

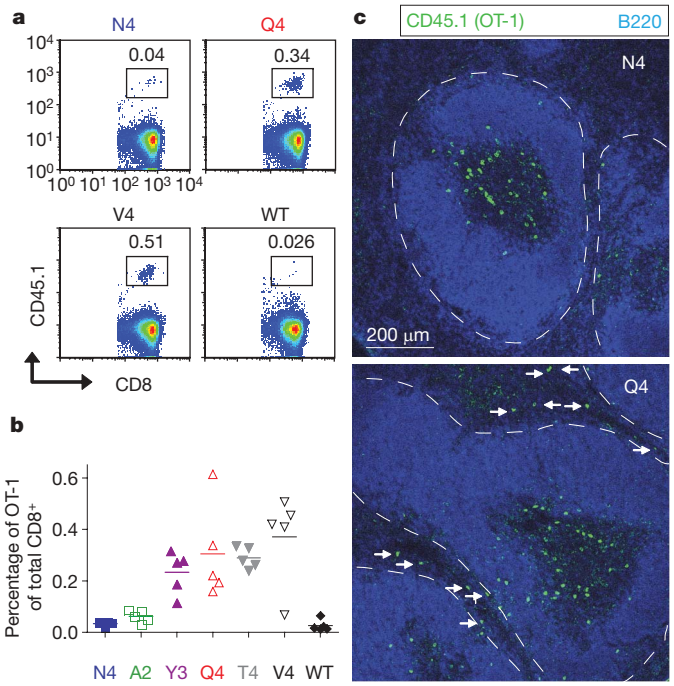


Figure 3 | The strength of TCR ligation determines migration kinetics. Ten thousand (**a**, **b**) and 3×10^3 (**c**) naive OT-1 cells were transferred into C57BL/6 mice, which were subsequently infected with Lm-N4ova or Lm-APLova strains. **a**, Representative flow plots of $CD8^+$-gated white blood cells showing the frequency of OT-1 cells at day 4 post-infection. **b**, Data for all APLs ($n = 5$). **c**, Splenic sections taken at day 4 post-infection show the distribution of OT-1 cells stimulated by N4 or Q4. OT-1 cells are green, B cells are blue. The border of white and red pulp is marked; arrows indicate OT-1 cells within the red pulp.

OT-1 cells previously exposed to any of the APLs expanded significantly (Fig. 4c). Control mice that initially received a wild-type *Listeria* infection contained far fewer OT-1 cells after VSV-N4ova challenge, indicating that all the APLs generated memory cells. These results were confirmed using $Rag2^{-/-}$ OT-1 cells (data not shown). When N4 and V4 memory OT-1 cells were transferred into new hosts, both expanded comparably after Lm-N4ova re-challenge. Notably, both memory populations also responded similarly after Lm-V4ova re-challenge, albeit the latter response was smaller than for Lm-N4ova (Supplementary Fig. 9). Thus, expansion in the recall response is determined by the strength of the recall stimulus and not by the priming antigen. These data, showing that even very weak TCR–ligand interaction drives the formation of functional memory T cells, are in line with reports indicating that lymphopaenia-driven homeostatic expansion, during which T cells also encounter only very weak TCR ligands, generates functional memory T cells[15].

The functional avidity of T-cell clones can change over the course of an immune response[16,17]. In our experiments, OT-1 T cells serve as an internal control for the changes in functional avidity at the clonal level. As such changes affect both endogenous and TCR transgenic T cells, the shift in the avidity of endogenous T cells compared with OT-1 T cells reflects changes in the T-cell repertoire. Our data therefore show that microbial challenge recruits $CD8^+$ T-cell clones with an extremely broad range of functional avidities for foreign antigen and that the composition of the initial repertoire rapidly changes to favour the 'best-fit', high-avidity T cells. This evolution of the repertoire is explained by the fact that very low-affinity TCR ligands are sufficient to induce rapid proliferation, but the length of the expansion period and the burst size directly correlate with the strength of TCR ligation so that cells stimulated by strong agonists go through more divisions than cells stimulated by weak agonists. Some viral pathogens can readily mutate their antigenic motifs[18,19]. The recruitment of a broad repertoire raises the possibility that some of the low-avidity effector and memory cells may be advantageous in controlling these pathogen variants[20].

TCR ligands of different affinity induce opposing cell fates during thymocyte development; that is, positive versus negative selection[21–23]. The peptides we incorporated into *Listeria* to study the response of peripheral OT-1 cells can be directly compared to the APLs used by Daniels *et al.* to study selection of immature OT-1 thymocytes[22]. The T4 ligand, used in both studies, is on the threshold

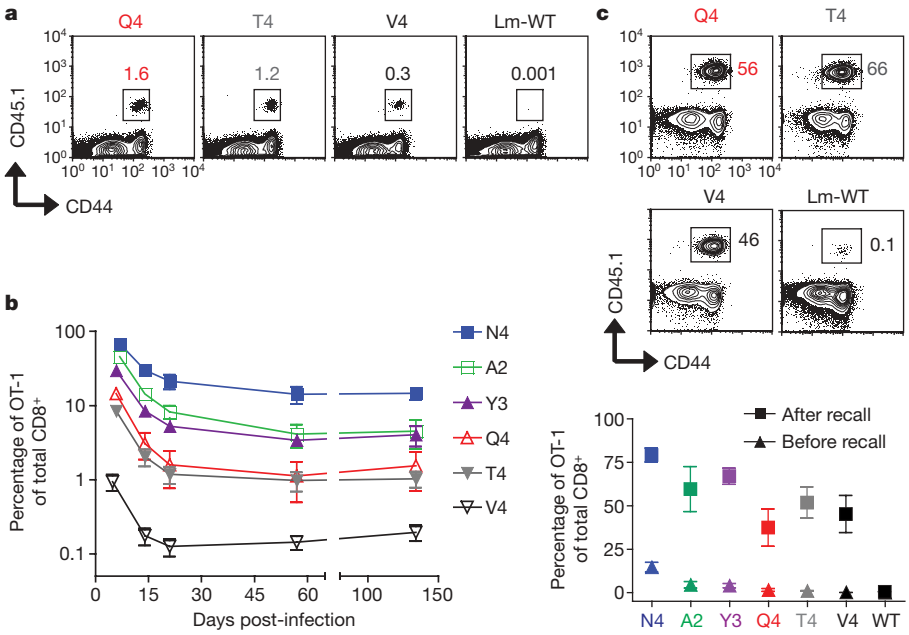


Figure 4 | Low-potency TCR ligands induce functional memory cells. Mice were grafted with congenic naive OT-1 cells and infected with Lm-WT, Lm-N4ova or Lm-APLova strains. **a**, Representative flow plots of the frequencies of memory OT-1 among blood $CD8^+$ T cells at day 138 post-infection. **b**, Contraction and memory kinetics for all mice and all APLs ($n = 5$) until day 138 post-infection. **c**, Mice were re-challenged on day 138 post-infection with VSV-N4ova, and the frequencies of OT-1 among blood $CD8^+$ T cells were determined 4 days later. Upper panels show representative flow plots and lower panels depict the data for all mice ($n = 5$); triangles indicate the frequency before, squares the frequency after, the recall. Bars show s.e.m.

©2009 Macmillan Publishers Limited. All rights reserved

of positive versus negative selection. The V4 ligand used here is much weaker than T4 (half-maximal effective concentration (EC_{50}) about 700 compared with about 70) and induces a much weaker peripheral OT-1 response than T4. Recombinant *Listeria* making the Q4H7 or Q7 ligands, which are closer in potency to T4 yet still capable of positive selection only, induced peripheral OT-1 expansion that was intermediate between T4 and V4 (Supplementary Fig. 10). These data raise the possibility that the minimum affinity needed to activate peripheral T cells during an infection overlaps the affinity required for thymocyte-positive selection on self-peptides, with implications for autoimmunity. Furthermore, although low- versus high-affinity ligands induce different fates in thymocytes[23], our data indicate that peripheral $CD8^+$ T cells do not show the same sharp ligand discrimination.

METHODS SUMMARY

CD45.1 or Thy1.1 congenic OT-1 TCR transgenic naive T cells were isolated using the CD8 Untouched Isolation Kit (Miltenyi Biotech) plus biotin-conjugated anti-CD44. Cells were transferred into C57BL/6 (Jackson Laboratories) mice one day before infection.

Recombinant *L. monocytogenes* strains were generated that stably expressed chicken ovalbumin ($AA_{134–387}$) containing either the native ligand $SIINFEKL_{257–264}$ or the APLs listed in Supplementary Fig. 2. The designation of the APLs indicates the substituted amino acid and the position within the SIINFEKL epitope. A previously described cassette[24] encoding the expression of secreted Ova was manipulated by site-directed mutagenesis to insert the APLs. The cassettes were cloned into a vector (pPL2) and stable *Listeria* recombinants were made as described[25]. Mice were infected intravenously with 1,000 colony-forming units of log-phase *Listeria*. For VSV-N4ova, mice were infected intravenously with 2×10^6 plaque-forming units.

Spleens were digested with Blendzyme 2 and DNase 1 (both Roche) for 1 h at 37 °C and mashed through a 100-μm cell strainer (Becton Dickinson) to obtain single-cell suspensions. Flow cytometry, intracellular cytokine staining and CFSE labelling were performed using standard procedures. For intracellular cytokine staining, the cells were first stimulated at 37 °C for 30 min with peptide, then 7 μM brefeldin was added and the incubation was continued for another 4.5 h. For CCR7 staining, cells were incubated for 1 h at 37 °C and then stained on ice with CCL19–IgG fusion supernatant[26]. To assess the level of non-specific binding of CCL19–IgG to OT-1, soluble CCL19 (R&D Systems) was added to control samples.

For immunofluorescent microscopy, spleens were fixed in 4% paraformaldehyde and kept in 30% sucrose overnight. Sections of 10–20 μM were cut and stained using standard procedures and analysed on a Zeiss 510 META confocal microscope.

Received 15 October; accepted 17 November 2008.
Published online 28 January 2009.

1. Antia, R., Ganusov, V. V. & Ahmed, R. The role of models in understanding $CD8^+$ T-cell memory. *Nature Rev. Immunol.* **5,** 101–111 (2005).
2. Blattman, J. N. *et al.* Estimating the precursor frequency of naive antigen-specific CD8 T cells. *J. Exp. Med.* **195,** 657–664 (2002).
3. Williams, M. A. & Bevan, M. J. Effector and memory CTL differentiation. *Annu. Rev. Immunol.* **25,** 171–192 (2007).
4. Turner, S. J., Doherty, P. C., McCluskey, J. & Rossjohn, J. Structural determinants of T-cell receptor bias in immunity. *Nature Rev. Immunol.* **6,** 883–894 (2006).
5. Wilson, D. B. *et al.* Specificity and degeneracy of T cells. *Mol. Immunol.* **40,** 1047–1055 (2004).
6. Badovinac, V. P., Haring, J. S. & Harty, J. T. Initial T cell receptor transgenic cell precursor frequency dictates critical aspects of the CD8(+) T cell response to infection. *Immunity* **26,** 827–841 (2007).
7. Henrickson, S. E. *et al.* T cell sensing of antigen dose governs interactive behavior with dendritic cells and sets a threshold for T cell activation. *Nature Immunol.* **9,** 282–291 (2008).
8. Joshi, N. S. *et al.* Inflammation directs memory precursor and short-lived effector CD8(+) T cell fates via the graded expression of T-bet transcription factor. *Immunity* **27,** 281–295 (2007).
9. Sarkar, S. *et al.* Functional and genomic profiling of effector CD8 T cell subsets with distinct memory fates. *J. Exp. Med.* **205,** 625–640 (2008).
10. Stemberger, C. *et al.* A single naive CD8+ T cell precursor can develop into diverse effector and memory subsets. *Immunity* **27,** 985–997 (2007).
11. Kaech, S. M. & Wherry, E. J. Heterogeneity and cell-fate decisions in effector and memory CD8+ T cell differentiation during viral infection. *Immunity* **27,** 393–405 (2007).
12. Gronski, M. A. *et al.* TCR affinity and negative regulation limit autoimmunity. *Nature Med.* **10,** 1234–1239 (2004).
13. Khanna, K. M., McNamara, J. T. & Lefrancois, L. *In situ* imaging of the endogenous CD8 T cell response to infection. *Science* **318,** 116–120 (2007).
14. Pamer, E. G. Immune responses to *Listeria monocytogenes*. *Nature Rev. Immunol.* **4,** 812–823 (2004).
15. Hamilton, S. E., Wolkers, M. C., Schoenberger, S. P. & Jameson, S. C. The generation of protective memory-like $CD8^+$ T cells during homeostatic proliferation requires $CD4^+$ T cells. *Nature Immunol.* **7,** 475–481 (2006).
16. Slifka, M. K. & Whitton, J. L. Functional avidity maturation of $CD8^+$ T cells without selection of higher affinity TCR. *Nature Immunol.* **2,** 711–717 (2001).
17. Xiao, Z., Mescher, M. F. & Jameson, S. C. Detuning CD8 T cells: down-regulation of CD8 expression, tetramer binding, and response during CTL activation. *J. Exp. Med.* **204,** 2667–2677 (2007).
18. Phillips, R. E. *et al.* Human immunodeficiency virus genetic variation that can escape cytotoxic T cell recognition. *Nature* **354,** 453–459 (1991).
19. Voeten, J. T. *et al.* Antigenic drift in the influenza A virus (H3N2) nucleoprotein and escape from recognition by cytotoxic T lymphocytes. *J. Virol.* **74,** 6800–6807 (2000).
20. Selin, L. K. *et al.* Memory of mice and men: CD8+ T-cell cross-reactivity and heterologous immunity. *Immunol. Rev.* **211,** 164–181 (2006).
21. Alam, S. M. *et al.* T-cell-receptor affinity and thymocyte positive selection. *Nature* **381,** 616–620 (1996).
22. Daniels, M. A. *et al.* Thymic selection threshold defined by compartmentalization of Ras/MAPK signalling. *Nature* **444,** 724–729 (2006).
23. Starr, T. K., Jameson, S. C. & Hogquist, K. A. Positive and negative selection of T cells. *Annu. Rev. Immunol.* **21,** 139–176 (2003).
24. Pope, C. *et al.* Organ-specific regulation of the CD8 T cell response to *Listeria monocytogenes* infection. *J. Immunol.* **166,** 3402–3409 (2001).
25. Lauer, P., Chow, M. Y., Loessner, M. J., Portnoy, D. A. & Calendar, R. Construction, characterization, and use of two *Listeria monocytogenes* site-specific phage integration vectors. *J. Bacteriol.* **184,** 4177–4186 (2002).
26. Hargreaves, D. C. *et al.* A coordinated change in chemokine responsiveness guides plasma cell movements. *J. Exp. Med.* **194,** 45–56 (2001).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank A. Gallegos for advice in making the recombinant *Listeria*, and X.-C. Pan, B. Dere, B. Paul and K. Rider for technical support. Financial support was received from the National Institutes of Health and the Howard Hughes Medical Institute (M.J.B.), and from the Juvenile Diabetes Research Foundation and the German Research Foundation (D.Z.).

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to M.J.B. (mbevan@u.washington.edu).

©2009 **Macmillan Publishers Limited. All rights reserved**

LETTERS

A single regulatory gene is sufficient to alter bacterial host range

Mark J. Mandel[1], Michael S. Wollenberg[1], Eric V. Stabb[2], Karen L. Visick[3] & Edward G. Ruby[1]

Microbial symbioses are essential for the normal development and growth of animals[1–3]. Often, symbionts must be acquired from the environment during each generation, and identification of the relevant symbiotic partner against a myriad of unwanted relationships is a formidable task[4]. Although examples of this specificity are well-documented, the genetic mechanisms governing it are poorly characterized[5]. Here we show that the two-component sensor kinase RscS is necessary and sufficient for conferring efficient colonization of *Euprymna scolopes* squid by bioluminescent *Vibrio fischeri* from the North Pacific Ocean. In the squid symbiont *V. fischeri* ES114, RscS controls light-organ colonization by inducing the Syp exopolysaccharide, a mediator of biofilm formation during initial infection. A genome-level comparison revealed that *rscS*, although present in squid symbionts, is absent from the fish symbiont *V. fischeri* MJ11. We found that heterologous expression of RscS in strain MJ11 conferred the ability to colonize *E. scolopes* in a manner comparable to that of natural squid isolates. Furthermore, phylogenetic analyses support an important role for *rscS* in the evolution of the squid symbiosis. Our results demonstrate that a regulatory gene can alter the host range of animal-associated bacteria. We show that, by encoding a regulator and not an effector that interacts directly with the host, a single gene can contribute to the evolution of host specificity by switching 'on' pre-existing capabilities for interaction with animal tissue.

Genomic technologies are facilitating major advances in understanding the relationships between metazoans and their bacterial symbionts. The analysis of unculturable endosymbionts has revealed complex genetic interdependence between host and bacteria amid patterns of genome reduction in endosymbiotic lineages[3]. Similarly, members of the human microbiota are being identified through metagenomic analysis[1,2], and the molecular communication between host and microbe has begun to be interpreted through transcriptional profiling[6]. Despite these advances, the mechanisms by which host–symbiont specificity develops in animal–bacterial interactions are not clear. Many animals, including humans, are born devoid of symbionts and must recruit their microbiota from the environment[7]. The process by which hosts and symbionts find each other to initiate a mutualism must be sensitive enough to identify the correct partner even when the symbiont is a minority constituent of the microbial community, and specific enough to exclude interlopers from gaining access to the host. The basis of species specificity is also poorly understood for pathogenic interactions, as similar congeneric bacteria often have distinct host ranges[8–10].

In this study we used a comparative genomics approach to reveal how bacteria–host specificity is established in the *E. scolopes*–*V. fischeri* mutualism. We took advantage of the fact that *V. fischeri* strain MJ11, which was isolated from the Japanese pinecone fish, *Monocentris japonica*[11], is unable to colonize *E. scolopes* efficiently. As such, comparison of MJ11 with natural squid symbionts provided a valuable system for examining the genomic basis of host specificity in an animal symbiont.

Although the genome sequence of squid-symbiotic *V. fischeri* strain ES114 is known[12,13], we determined here the sequence of the fish symbiont MJ11. Genome assembly of MJ11 was based on the ES114 model using a combination of PCR- and fosmid-based approaches. Genome sequencing also revealed a 179-kilobase (kb) circular plasmid in MJ11 that we term pMJ100, in which 82% of the open reading frames are annotated as hypothetical proteins, and which is distinct from the plasmid carried by ES114. Alignment of the assembled chromosomes revealed two circular MJ11 chromosomes that are co-linear to those in ES114 (Fig. 1a). Over 90% of ES114 open reading frames are shared by MJ11, and the orthologues have a median amino-acid identity of 98.8%. One exception to the high level of conservation was significant divergence observed specifically in the LuxR quorum-sensing system (Supplementary Fig. 1 and Supplementary Discussion).

Examination of ES114 genes for those that could facilitate specific recognition identified *rscS* as a promising candidate because its product acts during symbiotic initiation[14,15], and we discovered it to be absent in the MJ11 genome (Fig. 1b). RscS is a membrane-bound two-component sensor kinase that acts upstream of the response-regulator SypG[16]. SypG, a σ^{54}-dependent transcriptional activator, facilitates transcription of the 18-gene exopolysaccharide locus *sypA-R*[17]. Production of the Syp exopolysaccharide enables *V. fischeri* aggregation in squid-derived mucus during colonization of *E. scolopes*. During growth in culture, *syp* genes are expressed at low levels but can be induced by the plasmid-borne *rscS1* overexpression allele[15,17,18], leading to the production of robust biofilms. Because the MJ11 genome revealed an intact *syp* locus, we asked whether signal transduction downstream of *rscS* was maintained in MJ11 by introducing *rscS1*. As shown in Fig. 1c, d, *rscS1* in MJ11 induced multiple biofilm phenotypes, suggesting that the *syp* locus of MJ11 was functional. We therefore examined whether *rscS* was sufficient to allow MJ11 to colonize *E. scolopes* efficiently.

We tested ES114 and MJ11 for their ability to colonize aposymbiotic *E. scolopes* hatchlings in a 3-h inoculation assay. ES114 colonized successfully, whereas MJ11 failed to initiate colonization, even if present at a tenfold higher inoculum concentration (Fig. 1e). However, when provided with *rscS*$^+$ *in trans*, MJ11 was competent to colonize *E. scolopes* to levels comparable to those seen with the natural symbiont ES114 (Fig. 1e). Furthermore, the luminescence emitted by MJ11/*rscS*$^+$-colonized animals was 100-fold greater than that from animals colonized by ES114. The increased luminescence is consistent with that of the brighter fish symbiont[19], and was not influenced by plasmid carriage (Supplementary Fig. 2). This result argues that the carrying capacity of the juvenile squid light organ is specified by symbiont cell

[1]Department of Medical Microbiology and Immunology, University of Wisconsin School of Medicine and Public Health, 1550 Linden Drive, Madison, Wisconsin 53706, USA. [2]Department of Microbiology, University of Georgia, 828 Biological Sciences, Athens, Georgia 30602, USA. [3]Department of Microbiology and Immunology, Loyola University Chicago, Maywood, Illinois 60153, USA.

©2009 Macmillan Publishers Limited. All rights reserved

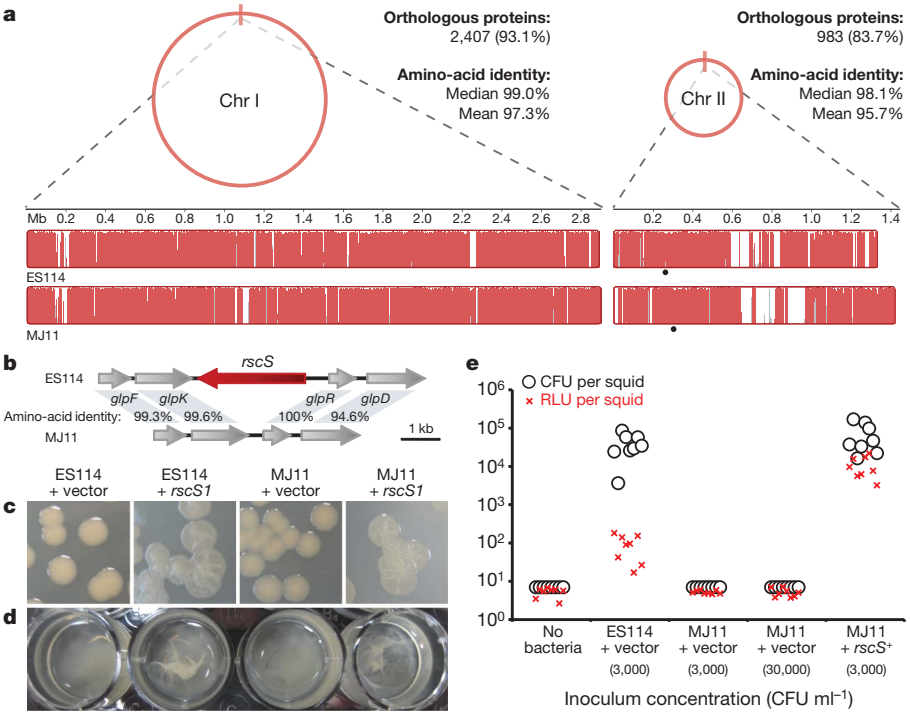


Figure 1 | *rscS* is sufficient to confer efficient colonization of *E. scolopes*. **a**, Mauve output shows each chromosome (chr) as one locally co-linear block. **b**, *rscS* is absent in MJ11 despite a high level of conservation in the surrounding locus. **c**, **d**, RscS-controlled biofilm phenotypes, including colony wrinkling (**c**) and pellicle formation (**d**) of strains harbouring either the vector control (pKV69) or the vector carrying *rscS1* (pKG11). **e**, Squid-colonization assay of strains harbouring either the vector control or the vector carrying $rscS^+$ (pLMS33), after 3 h inoculation and assayed at 48 h post-inoculation. Data points are individual animals. CFU, colony-forming units; RLU, relative light units.

number and not by the amount of luminescence emitted, provided that a minimum threshold of light production is achieved[20,21].

We next asked whether *rscS* was present in a collection of *V. fischeri* squid and fish isolates from the North Pacific Ocean to determine whether the gene's host distribution was consistent with a functional role in nature. All *V. fischeri* in the analysis revealed the presence of three representative *syp* genes (Fig. 2a). In contrast, although all of the squid isolates encoded *rscS*, regardless of geography, only five of the ten fish isolates encoded *rscS*; in four of these five the allele was significantly divergent. PCR amplification of the divergent alleles produced a band that was distinctively larger than the allele in ES114 and the rest of the squid isolates, whereas MJ12 was the only fish isolate that had this smaller squid–symbiont band (Fig. 2a).

We term the allele encoded by the smaller band $rscS_A$, and that encoded by the larger band $rscS_B$. $RscS_A$ was found in all assayed North Pacific squid isolates, and fish isolate MJ12, whereas $RscS_B$ was identified only in four fish isolates. Sequencing revealed that, within each type, the alleles for RscS are highly conserved (amino-acid identity at least 96%), but that divergence between the types was greater (84–86%; Supplementary Fig. 3a). The presence of an identical domain structure in all *V. fischeri* RscS proteins (Supplementary Fig. 3b) led us to ask whether there was detectable functional significance to this level of divergence.

All of the $rscS_A$ strains were competent to colonize *E. scolopes* squid efficiently (Fig. 2b). In contrast, strains lacking *rscS* or encoding the divergent $rscS_B$ were unable to colonize consistently. The defect appeared to be due to RscS function and not to the *syp* locus or other differences; introduction of $rscS^+$ from ES114 (A-type) into $rscS_B$-containing *mjapo*.8.1 conferred 100% colonization efficiency (Fig. 2b). The only fish-symbiotic strain that was competent to colonize *E. scolopes* reproducibly (MJ12) was also the only one with the conserved $rscS_A$ allele. Interruption of *rscS* in MJ12 abolished its ability to colonize *E. scolopes* (Fig. 2b), confirming that $rscS_A$ is both sufficient and necessary to colonize the squid host in these populations.

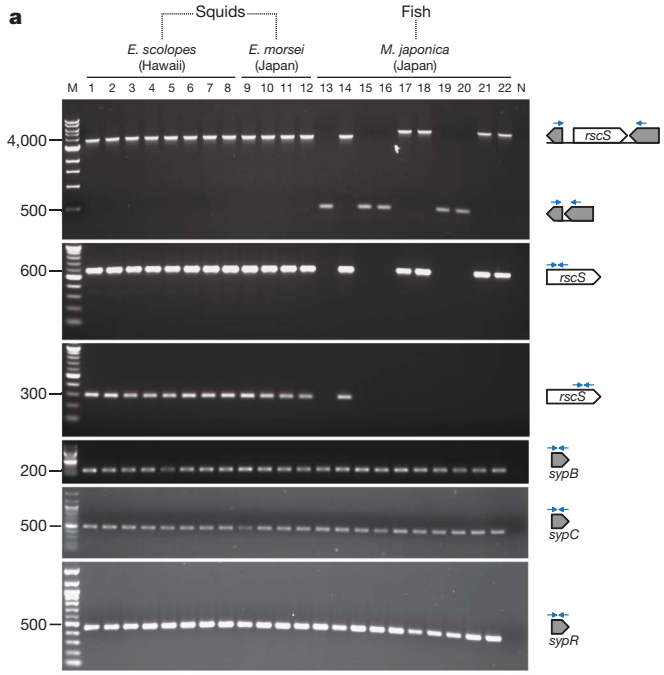
To understand the evolution of *rscS* and its role in determining specificity in nature, we reconstructed the phylogeny of *V. fischeri* strains (Fig. 3 and Supplementary Fig. 4) using three well-characterized loci. Strains encoding *rscS* formed a monophyletic group within *V. fischeri* that was statistically well supported. Parametric bootstrapping rejected the alternative hypothesis of non-monophyletic origin for the *rscS*-encoding strains, at a significance level of $P < 0.01$.

We propose a model for *rscS* evolution in the symbioses of North Pacific Ocean squids and fish. This model represents a parsimonious synthesis of the colonization and genomics data, within the phylogenetic framework. Specifically, we hypothesize that an acquisition event introduced *rscS* into the *V. fischeri* lineage before the expansion of this species into squid hosts in the North Pacific Ocean (Figs 3 and 4). An initial acquisition, followed by vertical transmission of *rscS* among *V. fischeri*, would predict both a similar guanine–cytosine (GC) content among all *rscS* alleles in the species, and a single conserved genomic location for the gene in all extant *V. fischeri* genomes. We confirmed these predictions, as the *rscS* alleles from *V. fischeri* have similar GC-content (Supplementary Fig. 3a) and are present in the same genomic position (Fig. 2a).

Because the fish isolates that contain *rscS* fall within the same clade as squid isolates (Fig. 3), we argue that the fish- and squid-symbiotic populations in Japan are, indeed, sympatric and that the *rscS*-containing fish isolates are descendents of squid symbionts. We hypothesize that $rscS_A$ diverged significantly in the fish host to generate the $rscS_B$ allele, which is not sufficient to allow these strains to colonize the squid host niche. Further, because $RscS_B$ maintained its reading frame and domain structure—despite significant amino-acid divergence and loss of function for squid colonization—we hypothesize that $RscS_B$ is fish-adapted, and may play a role in activating *syp*, and/or other targets, under fish-specific conditions. The identification of a fish isolate encoding $RscS_A$ provides strong evidence that the $rscS_A$ locus does not preclude successful fish colonization by *V. fischeri*, despite the low frequency of this allele among the fish isolates examined. Unfortunately, *M. japonica* eggs do not fully develop in the laboratory, so we are unable to test this aspect of our model by investigating *V. fischeri* colonization of fish[22].

There are two formal possibilities for how *rscS* first entered the *V. fischeri* lineage. Either a gene duplication/translocation event within *V. fischeri* led to the initial generation of *rscS*, or *rscS* was generated outside *V. fischeri* and then acquired by horizontal gene transfer. We have found no DNA sequences paralogous to either of the *rscS* alleles in the full genomes of ES114 or MJ11. These data are also the most compelling, if indirect, evidence supporting the proposal of horizontal gene transfer. Nonetheless, it is difficult to reconstruct the event that introduced *rscS* to *V. fischeri* based on the usual criteria that define larger genomic islands (for example, direct repeats or insertion elements in flanking DNA, or aberrant codon usage or GC content within

©2009 Macmillan Publishers Limited. All rights reserved



b

Lane	Strain	Host	*rscS* type	Number of *E. scolopes* colonized/attempted
1	ES114	*Es*	A	29/29 0/8 ...[*rscS*$^-$]
2	KB1A97	*Es*	A	16/16 0/8 ...[*rscS*$^-$]
3	KB2B1	*Es*	A	8/8
4	KB5A1	*Es*	A	8/8
5	ES213	*Es*	A	8/8
6	MB11B1	*Es*	A	8/8
7	MB13B1	*Es*	A	8/8
8	MB14A1	*Es*	A	8/8
9	EM17	*Em*	A	8/8
10	EM18	*Em*	A	8/8
11	EM24	*Em*	A	8/8
12	EM30	*Em*	A	8/8
13	MJ11	*Mj*	–	0/21
14	MJ12	*Mj*	A	20/20 0/8 ...[*rscS*$^-$]
15	*mjapo*.2.1	*Mj*	–	1/8
16	*mjapo*.3.1	*Mj*	–	0/8
17	*mjapo*.4.1	*Mj*	B	0/8
18	*mjapo*.5.1	*Mj*	B	0/8
19	*mjapo*.6.1	*Mj*	–	1/8
20	*mjapo*.7.1	*Mj*	–	0/8
21	*mjapo*.8.1	*Mj*	B	0/8 4/4 ...[+*rscS*$_A$]
22	*mjapo*.9.1	*Mj*	B	1/8

Figure 2 | The presence of *rscS* is correlated to a natural association with squid and to the ability to colonize *E. scolopes* experimentally. **a**, PCR assay for the presence of *rscS* and *syp* genes with primer sets (top to bottom) *rscS*-flank, *rscS*-internal1, *rscS*-internal2, *sypB*-internal, *sypC*-internal or *sypR*-internal. M, markers; N, negative control. **b**, Colonization competence of strains in **a**, colonized as described in Fig. 1, at a concentration of approximately 5×10^3 colony-forming units per millilitre. Entries in red indicate *rscS*$^-$ derivatives; entry in blue indicates a strain carrying a plasmid (pLMS33) encoding *rscS*$^+$ from ES114. *Em*, *Euprymna morsei*; *Es*, *E. scolopes*; *Mj*, *M. japonica*.

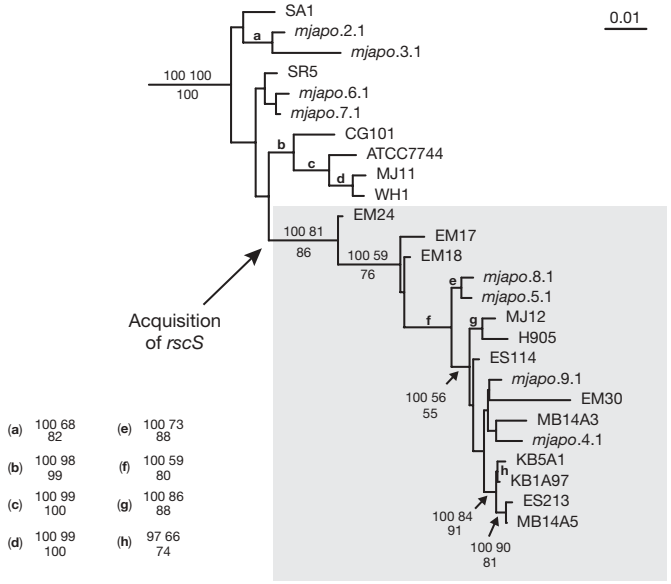


Figure 3 | *rscS* entered the *V. fischeri* lineage before colonization of squid in the North Pacific Ocean. Bayesian phylogeny of 26 *V. fischeri* strains using the concatenation of three loci (*recA*, *mdh*, *katA*). The tree was rooted with *Aeromonas salmonicida* subsp. *salmonicida* A449. Support for branches is indicated in cases where support is greater than 50% for all three: Bayesian posterior probabilities (top left), maximum likelihood bootstrap (top right), and maximum parsimony bootstrap (bottom). *rscS* DNA was detected in all strains in the grey box and in none of the strains outside the box. Bar: 0.01 expected changes per site.

rscS)[23]. Furthermore, the only convincing orthologue of *rscS* outside *V. fischeri* is *V. shiloi* AK1 VSAK1_16757. If a horizontal transfer event were responsible for *rscS* transmission into *V. fischeri*, the *V. shiloi* orthologue is unlikely to be the source: the GC content of ES114 *rscS* is 31.7%, or 6.6% below the ES114 genome average (38.3%). The *V. shiloi* orthologue has a GC content (41.0%) that is even higher than this average, and 9.3% higher than that of the ES114 *rscS* allele.

Attempts to understand the molecular basis of host specificity have been unsuccessful in many pathogen–host animal interactions. *Salmonella enterica* serovar Typhi can infect only humans, whereas serovar Typhimurium has a broad host range that includes mice[9]. Although the conserved regions of the genomes of these two strains are over 97% identical, efforts to account for this differential host specificity have not succeeded. Similarly, different *Brucella* species share over 98% identity across 90% of their genes, yet exhibit strict host specificity; the molecular basis of this specificity remains unclear[10]. In contrast, the study of mutualisms is providing insight into how specificity develops. In plant-associated bacteria, work from many laboratories has established nitrogen-fixing, nodulating rhizobia as the best-understood system for the development and evolution of host specificity[24]. Bacteria secrete Nod factors—lipo-chitooligosaccharide signals—to the plant host, and host-strain-specific backbone modifications encoded by the bacteria lead to relationship specificity. Recently, in an animal–bacterial mutualism, the *nilABC* genes of *Xenorhabdus nematophila* were characterized as sufficient for colonization of *Steinernema carpocapsae* worms by congeneric *Xenorhabdus* bacteria[25].

In contrast to rhizobia and *Xenorhabdus*, in which specificity comes either from the modification of a secreted signal or from structural proteins in the cell envelope, respectively, *rscS*-mediated specificity in *V. fischeri* is novel because the immediate effect of cytoplasmic RscS is on bacterial gene expression, which only subsequently has an effect on the interaction with the host. Because RscS is a signal transduction protein, the evolutionary consequence of the introduction of *rscS* appears to be a re-programming of inherent *V. fischeri* capabilities to expand the host range into squid populations that *V. fischeri* could not previously colonize, or could colonize only inefficiently. It remains a mystery as to why the *syp* genes are conserved in *V. fischeri* strains that are naive to *rscS* (for example, MJ11). That such *syp* clusters are functional, and ancestral to *rscS* in *V. fischeri*, strongly suggests that regulation of *syp* in these strains may be achieved in a manner independent of *rscS*. In support of this, there are *V. fischeri* isolated from the Mediterranean Sea that lack *rscS*, yet have *syp* genes and colonize squid hosts of a different genus through morphological structures that are conserved with those of *E. scolopes*[26].

Our study indicates that a regulatory gene is sufficient to alter host range in an animal–bacterial mutualism. The fundamental biological question of how animal–bacterial partnerships are established has been difficult to access through investigations of pathogenic interactions. In

©2009 Macmillan Publishers Limited. All rights reserved

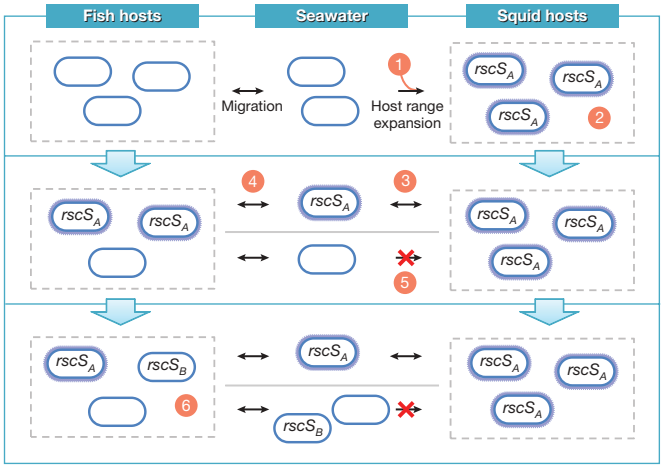


Figure 4 | Reconstruction of the evolution of *V. fischeri* symbioses in the North Pacific Ocean. All strains are genotypically *syp*$^+$; strains that are phenotypically Syp$^+$ in squid are illustrated with the grey mottled outline. The model reconstructs host range expansion into squid by (1) acquisition of *rscS*$_A$, and (2) subsequent developmentally appropriate expression of the ancestral *syp* polysaccharide genes. (3) Gene flow into planktonic strains and (4) into fish symbionts accounts for the presence of *rscS*$_A$ in fish isolates, but (5) *rscS*$^-$ strains maintain a colonization incompatibility for squid. (6) We postulate that *rscS*$_B$ evolved from *rscS*$_A$ in fish.

contrast, mutualism evolves to confer joint benefits to its partners, and relies on a strict specificity for this outcome: that is, entry of only a few appropriate symbionts and exclusion of the many non-specific interlopers. The evolution of developmental mechanisms to winnow the appropriate partner(s) is a hallmark of all horizontally acquired mutualisms. The binary squid–*Vibrio* system thus represents a valuable model in which to interrogate the mechanisms that underlie the development of bacteria–host specificity.

METHODS SUMMARY

The previously deposited draft genome of *V. fischeri* MJ11 was assembled into the final scaffold by comparing contigs to the ES114 assembly using Mauve[27]. Hypotheses were tested by PCR across the contig gaps or by sequencing of fosmids spanning the gap. In a few cases, no PCR product was produced, and the model was refined by rearranging contigs and retesting. In this manner, all of the contigs were arranged relative to ES114, and contig gaps that could be spanned by PCR were sequenced to complete the gap sequence. Three gaps on chromosome I contained tandem (at least two) *rrn* operons; in these cases, the sequence flanking the gap was PCR-amplified through the first ribosomal RNA (rRNA) gene at each end of the *rrn* array so that the completed genome was expected to contain all predicted open reading frames in *V. fischeri* MJ11.

We identified and corrected frameshift and nonsense mutations in the genome model[12], and the final sequences were annotated by the J. Craig Venter Institute (JCVI). To identify ES114–MJ11 orthologues, we performed reciprocal BLASTP searches between the predicted proteomes. Percentage identity was used to score results, and at least 60% coverage of each protein was demanded.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 18 August; accepted 20 November 2008.
Published online 1 February 2009.

1. Aas, J. A., Paster, B. J., Stokes, L. N., Olsen, I. & Dewhirst, F. E. Defining the normal bacterial flora of the oral cavity. *J. Clin. Microbiol.* **43,** 5721–5732 (2005).
2. Gill, S. R. *et al.* Metagenomic analysis of the human distal gut microbiome. *Science* **312,** 1355–1359 (2006).
3. Moran, N. A. Symbiosis as an adaptive process and source of phenotypic complexity. *Proc. Natl Acad. Sci. USA* **104** (suppl. 1), 8627–8633 (2007).
4. Mazmanian, S. K., Round, J. L. & Kasper, D. L. A microbial symbiosis factor prevents intestinal inflammatory disease. *Nature* **453,** 620–625 (2008).
5. Groisman, E. A. & Casadesús, J. The origin and evolution of human pathogens. *Mol. Microbiol.* **56,** 1–7 (2005).
6. Sonnenburg, J. L. *et al.* Glycan foraging *in vivo* by an intestine-adapted bacterial symbiont. *Science* **307,** 1955–1959 (2005).
7. Xu, J. & Gordon, J. I. Honor thy symbionts. *Proc. Natl Acad. Sci. USA* **100,** 10452–10459 (2003).
8. Brinig, M. M., Register, K. B., Ackermann, M. R. & Relman, D. A. Genomic features of *Bordetella parapertussis* clades with distinct host species specificity. *Genome Biol.* **7,** R81 (2006).
9. Edwards, R. A., Olsen, G. J. & Maloy, S. R. Comparative genomics of closely related salmonellae. *Trends Microbiol.* **10,** 94–99 (2002).
10. Rajashekara, G., Glasner, J. D., Glover, D. A. & Splitter, G. A. Comparative whole-genome hybridization reveals genomic islands in *Brucella* species. *J. Bacteriol.* **186,** 5040–5051 (2004).
11. Ruby, E. G. & Nealson, K. H. Symbiotic association of *Photobacterium fischeri* with the marine luminous fish *Monocentris japonica*; a model of symbiosis based on bacterial studies. *Biol. Bull.* **151,** 574–586 (1976).
12. Mandel, M. J., Stabb, E. V. & Ruby, E. G. Comparative genomics-based investigation of resequencing targets in *Vibrio fischeri*: focus on point miscalls and artefactual expansions. *BMC Genomics* **9,** 138 (2008).
13. Ruby, E. G. *et al.* Complete genome sequence of *Vibrio fischeri*: a symbiotic bacterium with pathogenic congeners. *Proc. Natl Acad. Sci. USA* **102,** 3004–3009 (2005).
14. Visick, K. L. & Skoufos, L. M. Two-component sensor required for normal symbiotic colonization of *Euprymna scolopes* by *Vibrio fischeri*. *J. Bacteriol.* **183,** 835–842 (2001).
15. Yip, E. S., Geszvain, K., DeLoney-Marino, C. R. & Visick, K. L. The symbiosis regulator *rscS* controls the *syp* gene locus, biofilm formation and symbiotic aggregation by *Vibrio fischeri*. *Mol. Microbiol.* **62,** 1586–1600 (2006).
16. Hussa, E. A., Darnell, C. L. & Visick, K. L. RscS functions upstream of SypG to control the *syp* locus and biofilm formation in *Vibrio fischeri*. *J. Bacteriol.* **190,** 4576–4583 (2008).
17. Yip, E. S., Grublesky, B. T., Hussa, E. A. & Visick, K. L. A novel, conserved cluster of genes promotes symbiotic colonization and σ^{54}-dependent biofilm formation by *Vibrio fischeri*. *Mol. Microbiol.* **57,** 1485–1498 (2005).
18. Geszvain, K. & Visick, K. L. Multiple factors contribute to keeping levels of the symbiosis regulator RscS low. *FEMS Microbiol. Lett.* **285,** 33–39 (2008).
19. Stabb, E. V., Schaefer, A., Bose, J. L. & Ruby, E. G. in *Chemical Communication among Bacteria* (eds Winans, S. C. & Bassler, B. L.) 233–250 (ASM Press, 2008).
20. Bose, J. L., Rosenberg, C. S. & Stabb, E. V. Effects of *luxCDABEG* induction in *Vibrio fischeri*: enhancement of symbiotic colonization and conditional attenuation of growth in culture. *Arch. Microbiol.* **190,** 169–183 (2008).
21. Visick, K. L., Foster, J., Doino, J., McFall-Ngai, M. & Ruby, E. G. *Vibrio fischeri lux* genes play an important role in colonization and development of the host light organ. *J. Bacteriol.* **182,** 4578–4586 (2000).
22. Haygood, M. G. Light organ symbioses in fishes. *Crit. Rev. Microbiol.* **19,** 191–216 (1993).
23. Hacker, J. & Kaper, J. B. Pathogenicity islands and the evolution of microbes. *Annu. Rev. Microbiol.* **54,** 641–679 (2000).
24. Long, S. R. Genes and signals in the rhizobium–legume symbiosis. *Plant Physiol.* **125,** 69–72 (2001).
25. Cowles, C. E. & Goodrich-Blair, H. The *Xenorhabdus nematophila nilABC* genes confer the ability of *Xenorhabdus* spp. to colonize *Steinernema carpocapsae* nematodes. *J. Bacteriol.* **190,** 4121–4128 (2008).
26. Foster, J. S., Von Boletzky, S. & McFall-Ngai, M. J. A comparison of the light organ development of *Sepiola robusta* Naef and *Euprymna scolopes* Berry (Cephalopoda: Sepiolidae). *Bull. Mar. Sci.* **70,** 141–153 (2002).
27. Darling, A. C., Mau, B., Blattner, F. R. & Perna, N. T. Mauve: multiple alignment of conserved genomic sequence with rearrangements. *Genome Res.* **14,** 1394–1403 (2004).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank: P. Dunlap for sharing bacterial strains; N. Perna, J. Glasner, K. Geszvain, D. Baum, J. Johnson, M. Sarmiento and S. Ferriera for technical assistance; A. Wier, N. Bekiares, R. Gates and the Hawaii Institute of Marine Biology for animal facilities and care; J. McCosker and the Steinhart Aquarium for access to fish specimens; M. McFall-Ngai, H. Goodrich-Blair, C. Brennan and J. Troll for discussions; and L. Proctor for project support. MJ11 genome sequencing was funded by the Gordon and Betty Moore Foundation Marine Microbial Genome Sequencing Project; E.G.R. and co-workers are funded by the National Institutes of Health–National Center for Research Resources and the National Science Foundation Division of Integrative Organismal Systems; E.V.S. is funded by a National Science Foundation CAREER Award; K.L.V. is funded by the National Institute of General Medical Sciences; M.J.M. is funded by a National Institute of General Medical Sciences National Research Service Award Postdoctoral Fellowship; M.S.W. is funded by a National Science Foundation Predoctoral Fellowship and a National Institutes of Health Molecular Biosciences Training Grant to the University of Wisconsin.

Author Contributions M.J.M. designed the experiments, performed all work not described below, and wrote the paper. M.S.W. conducted the phylogenetic studies. K.L.V. constructed plasmids and strains, and imaged biofilm phenotypes. M.J.M. planned and performed the genome assembly and analytics, and M.J.M., E.V.S. and E.G.R. analysed the bioinformatics results.

Author Information The *recA*, *mdh*, *katA* and *rscS* sequence data from the additional strains described in the article are deposited in GenBank under accession numbers EU907941–EU908017; MJ11 genome data are deposited under accession numbers CP001133, CP001134 and CP001139. Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to M.J.M. (mmandel@wisc.edu).

©2009 Macmillan Publishers Limited. All rights reserved

METHODS

Bacterial growth, strains and plasmids. Standard microbial techniques were used to construct strains and plasmids[28]. Growth of *V. fischeri* was at 20–28 °C with aeration. Media for growth of *V. fischeri* was LBS[29], and for *E. coli* was LB[30] or brain heart infusion (Bacto, Becton Dickinson). Antibiotics used included: chloramphenicol (2.5 μg ml^{-1} for *V. fischeri*, 25 μg ml^{-1} for *E. coli*), erythromycin (5 μg ml^{-1} for *V. fischeri*, 150 μg ml^{-1} for *E. coli*), and tetracycline (5 μg ml^{-1} for *V. fischeri*).

Plasmids pKV69, pLMS33 and pKG11 were introduced into *V. fischeri* by triparental conjugation as described previously[28] with *E. coli* carrying the pEVS104 helper plasmid. Briefly, overnight cultures of the following strains were used for the reaction. One hundred microlitres from each of donor and helper *E. coli* strains were pelleted at 16,000*g* for 2 min in a microfuge tube, and the supernatant was aspirated. One hundred microlitres of recipient *V. fischeri* was added to the same tube, and pelleted as above. After aspiration of the supernatant, the pellet was re-suspended in 10 μl LBS, and the entire 10 μl was spotted onto an LBS agar plate, and incubated at 28 °C overnight. The spot was re-suspended in 500 μl LBS, and 50 μl were plated onto selective media (LBS–chloramphenicol) to select for plasmid transfer.

The *rscS* mutagenesis plasmid pKV188 was constructed by subcloning an approximately 700-base-pair (bp) internal *rscS* fragment from ES114, terminating in the internal *Pst*I site, into the *Kpn*I/*Pst*I sites of pEVS122 (ref. 28). Mutagenesis of *rscS* in the strains noted (Supplementary Table 1) was as described previously[31]; briefly, after triparental conjugation as above, integration of the suicide vector was identified by selection of the entire mating spot on LBS–erythromycin.

V. fischeri strain MJ11 (alias MJ101) was isolated by sterile expression of the light-organ sample from a live *M. japonica* at the Steinhart Aquarium, in February 1991. *M. japonica* symbiont strains denoted '*mjapo.#.#*' were shared by P. Dunlap.

Sources of other strains were as noted in Supplementary Table 1.

MJ11 genome assembly. The *V. fischeri* MJ11 draft genome was sequenced by the JCVI as part of the Gordon and Betty Moore Foundation Marine Microbial Genome Sequencing Project. Cloning and shotgun sequencing was performed at JCVI; draft coverage was obtained at 8.57-fold coverage, and the contigs were previously deposited into GenBank as a whole genome shotgun sequencing project, accession number ABIH00000000, project version 01.

Step 1: sealing contig gaps. *V. fischeri* MJ11 contigs from the above project were aligned to the ES114 genome using Mauve[27] and Projector2 (ref. 32). Alignment of all contigs that were from multiple reads (greater than 2 kb), excluding repetitive rDNA sequences, identified strong matches to the ES114 chromosome I and chromsome II, with the exception of 179-kb contig number 1101159000798. Owing to the high level of conservation between the strains, the corresponding ES114 sequences at all contig gaps were used as estimates of the gap lengths—and as guides for sequencing primers across long gaps—and primers were designed to amplify across each contig-gap, including at least an additional 200 bp of overlap with each adjacent contig, and extending beyond regions of repetitive DNA at either side of the gap boundary. PCR primers were designed with Primer3Plus[33] and sequencing primers were designed at the SGD website (http://seq.yeastgenome.org/cgi-bin/web-primer). In some cases, custom software was used to assist in the identification of probable-unique regions for primer-binding sites within extended repeat regions.

Contig number 1101159000798 was unique in bearing no homology to ES114 or to any extended sequence in GenBank at the start of this project (autumn 2006). We postulated that this represented a large (circular) plasmid in MJ11, and consistent with this hypothesis primers pointing outward from both ends of the contig together amplified a small fragment (less than 1 kb).

Step 2: PCR-walking into tandem rRNA operons. Three of the contig gaps contained tandem rRNA operons, and PCR across the entire gap was unsuccessful. We initiated PCR-walking into each gap by amplifying from one end of the gap to a conserved region that was distal to the 16S–23S spacer on that end of the gap, using primers listed in Supplementary Table 2. Using this approach, we identified the six unique spacer sequences that comprised the terminal rRNA spacers for the three gaps, leaving only additional rRNA (and potentially transfer RNA (tRNA)) sequences remaining to be sealed within each of these three regions. These contigs were still assembled into a scaffold and submitted for deposition so that information about the relative positions of DNA sequences within the molecule (chromosome I) was preserved.

Step 3: re-sequencing of select targets. We previously published methodologies to identify and correct sequencing errors in microbial genomes[12]. We applied this technology to the MJ11 genome and identified 15 high-priority re-sequencing targets. Nine of these sites were in fact in error, and we corrected these errors for inclusion in the final genome release. The PCR/sequencing primers used to target these regions are listed in Supplementary Table 2.

After assembly, the sequence was resubmitted to the JCVI Annotation Service and post-processing at JCVI and the National Center for Biotechnology Information, and deposited into GenBank as listed in Supplementary Table 3.

Orthologue comparisons between ES114 and MJ11. We compared the predicted proteomes from both ES114 chromosomes against both MJ11 chromosomes. Reciprocal exhaustive BLASTP[34] searches were performed with an expect cutoff of 10. Results were filtered to demand that the query length and subject length each be a minimum of 60% of their respective total lengths. Among the remaining results for each query protein, best-hits were scored by percentage amino-acid identity, and additional results were included for analysis if they scored at least 70% of the maximum score for that query. ES114–MJ11 protein pairs included on reciprocal lists were candidate orthologues, and for pairs in which there was a duplicate of query or subject protein, manual assignment of orthology was curated using the parameters of percentage amino-acid identity, percentage of each protein aligned and the local genomic context (synteny) of the two proteins.

Biofilm phenotypes. The two biofilm phenotypes evaluated were colony morphology and pellicle formation. To assay for the ability to form wrinkled colonies, cells were streaked onto LBS–tetracycline agar and the plates were incubated at room temperature for two days. To assay for the ability to form pellicles, cells were inoculated into HEPES minimal medium[35] containing 0.3% casamino acids, 0.2% glucose and tetracycline at a final concentration of 30 μg ml^{-1}. After overnight growth at 28 °C with shaking, cells were diluted to an $OD_{600\,nm}$ of 0.1 in fresh HEPES minimal medium. Three millilitres of cell suspensions were introduced into the wells of a 12-well microtitre dish, and the cells were incubated statically at room temperature for five days. To facilitate visualization and imaging of the pellicles, the media surface was disturbed with a pipette tip, resulting in clumps of cells if a pellicle had formed.

Squid colonization assays. Juvenile *E. scolopes* hatchlings were collected aposymbiotically, and washed in Instant Ocean (Aquarium Systems) that was filter-sterilized through a 22-μm pore-sized Nalgene filter (FSIO: filter-sterilized Instant Ocean). Overnight cultures of bacteria in LBS were subcultured 1:40, grown for 70 min, assayed for $OD_{600\,nm}$, and then inoculated into 40 ml FSIO-squid at a volume equivalent to 1.25 μl per $OD_{600\,nm}$. The inoculum was plated onto LBS plates to confirm that the bacterial concentration was 2×10^3 to 10×10^3 colony-forming units per millilitre. Squid were washed with fresh FSIO at 3 and 24 h post-inoculation. Individual animal luminescence was recorded at 48 h post-inoculation before they were euthanized at 48 h by freezing at −80 °C. Symbiont colony-forming units per squid were determined by homogenizing thawed animals, and plating the homogenates onto LBS. For experiments involving pKV69-series plasmids, squid hatchlings were maintained in FSIO containing chloramphenicol at a final concentration of 2.5 μg ml^{-1}.

Luminescence of animals is reported as RLU (1 RLU $\approx 1.98 \times 10^4$ quanta per second) $= 24 \times lum$, where *lum* is the recorded luminescence of a single animal in a TD20/20 luminometer, with recordings performed in 4 ml FSIO in glass scintillation vials at 51.9% sensitivity with integration for 6 s. Animals with RLU > 25 were scored as colonized. Because the strains that failed to colonize *E. scolopes* were significantly brighter than ES114, this metric served as a conservative measure of colonization competency for the set of isolates examined in this study.

PCR amplification for MJ11 genome closure. PCR amplification was conducted using Platinum Taq DNA Polymerase High-Fidelity (Invitrogen). Fifty-microlitre reactions contained: 50 ng MJ11 genomic DNA, 1× reaction buffer, 0.2 mM of each dNTP, 2 mM $MgSO_4$, 0.25 μM of each primer, and 1 U DNA polymerase. At least three independent PCR reactions were combined for sequencing to minimize the effect of PCR error. Thermal cycling was conducted in a PTC-200 thermal cycler (MJ Research): 95 °C for 2 min; then 30 cycles of 95 °C for 30 s, 55 °C for 30 s, 68 °C for 30 s to 1 min per kilobase amplified; then 68 °C for 5 min.

Products greater than 5 kb were amplified with Platinum Pfx DNA Polymerase (Invitrogen) before sequencing. Fifty-microlitre reactions contained: 50 ng MJ11 genomic DNA, 2× reaction buffer, 0.3 mM of each dNTP, 1 mM $MgSO_4$, 0.30 μM of each primer, and 1 U DNA polymerase. Thermal cycling was conducted as above.

Primers for MJ11 genome closure are listed in Supplementary Table 2.

Diagnostic PCR amplification of rscS and syp genes. Conditions were as described above (MJ11 genome closure) with the following alterations. Template preparation consisted of bacterial strains grown overnight in LBS, diluted 1:100 in dH_2O, and then used as template in the PCR reactions at a dilution of 1:10. Annealing temperature for the *sypR*-internal primer set was 50 °C. Primers for diagnostic amplification of *rscS* and *syp* locus genes are listed in Supplementary Table 2.

PCR amplification for phylogenetic analyses. PCR amplification was conducted using GoTaq (Promega). Bacterial strains were grown overnight in

©2009 Macmillan Publishers Limited. All rights reserved

LBS, diluted 1:100 in dH_2O, and then used as template in the PCR reactions at a dilution of 1:10. Twenty-five-microlitre reactions contained: template preparation (2.5 μl), 1 × colourless reaction buffer, 0.2 mM of each dNTP, 0.9 μM of each primer and 1 U DNA polymerase. Thermal cycling was conducted in a PTC-200 thermal cycler (MJ Research): 95 °C for 3 min; then 26 cycles of 94 °C for 30 s, 55 °C for 30 s, 72 °C for 1 min; then 72 °C for 10 min. Primers for phylogenetic analyses are listed in Supplementary Table 2.

DNA sequencing. Sanger-type sequencing of PCR products for MJ11 genome assembly, and for phylogenetic analyses, was performed at the University of Washington High-Throughput Genomics Unit (Seattle, Washington) and the University of Wisconsin Biotechnology Center DNA Sequencing Facility (Madison, Wisconsin), with the primers listed in Supplementary Table 2.

Sequence data for *rscS* from *V. fischeri* ES114 and from *V. shiloi* AK1 are from GenBank accession numbers AF319618 and EDL55668, respectively. Sequence data for phylogenetic analysis of the following strains are as noted: *A. salmonicida* subsp. *salmonicida* (GenBank accession number CP000644), *Vibrio harveyi* BB120 (CP000789, CP000790), *V. parahaemolyticus* RIMD2210633 (BA000031, BA000032), *Photobacterium profundum* 3TCK (AAPH00000000) and *V. fischeri* ES114 (CP000020, CP000021).

Phylogenetic analyses. Sequences from the three loci (*recA*, *mdh*, *katA*) were aligned using ClustalX 1.83 (ref. 36), and trimmed and concatenated using custom Perl scripts (https://mywebspace.wisc.edu/wollenberg/web/science/scripts/scripts.html) and MEGA4 (ref. 37). The best-fit model of DNA substitution and parameter estimates used for tree reconstruction was chosen by performing hierarchical likelihood ratio tests on these data, as implemented in PAUP* 4.0b10 (ref. 38) and MODELTEST 3.7 (ref. 39); this model was SYM + I + G, a submodel under the GTR + I + G (general time reversible with gamma-rate distribution across sites and a proportion of invariant sites). Phylogenetic trees' clade topology and confidence were studied using three approaches. (1) MrBayes 3.1 (ref. 40), implementing the Markov chain Monte Carlo method with an evolutionary model set to GTR with gamma-distributed rate variation across sites and a proportion of invariable sites, was run for 5,000,000 generations using the CIPRES project portal (http://www.phylo.org/). Sample frequency was 1,000, creating a posterior probability distribution of 5,000 trees; when summarizing the substitution model parameters and trees, 1,250 trees were discarded as burn-in to address potential chain instability. (2) Maximum likelihood analyses were performed using the genetic algorithm approach of GARLI[41] as implemented in the CIPRES portal, with an evolutionary model set to GTR with gamma-distributed rate variation across sites, and a proportion of invariable sites. Bootstrap analysis of 1,000 replications was used to assess the support for internal nodes. (3) Unweighted maximum parsimony analysis and bootstrap were performed by PAUP* (1,000 replications) using heuristic searches implementing tree bisection and reconnection branch-swapping to find the shortest trees and assess the support for internal nodes. For maximum likelihood and Bayesian approaches, the process was independently repeated three times to ensure arrival at a similar, most-likely tree topology. Resulting trees were rooted with *A. salmonicida* subsp. *salmonicida* A449 as the outgroup (Supplementary Fig. 4).

Parametric bootstrap analysis. From our original data, the difference in likelihood scores between an unconstrained phylogeny and a constrained phylogeny with a non-monophyletic, *rscS*-containing clade was calculated. One hundred simulations of the data set were created using the constrained topology; likelihood scores were produced from these 100 simulated data sets both with and without the constraint of the non-monophyly of the *rscS*-containing clade using the software PAUP*. Our null hypothesis of the significance of the constraint of non-monophyly of the *rscS*-containing clade within our initial phylogeny was rejected based on analysis of the resulting likelihood ratio distribution ($P < 0.01$).

Codon usage. Whole genome codon usage for ES114 was analysed by the methods of Karlin and Mrázek[42], against a set of reference sequences from ES114, which included ribosomal proteins, chaperones and transcription/translation factors. Calculations were performed on the Computational Microbiology Laboratory server (http://www.cmbl.uga.edu/software/phxpa.html), and the codon usage of *rscS* was predicted to be neither highly expressed (PHX) nor alien (PA). All comparisons of codon bias performed placed *rscS* in the 95% confidence interval for the ES114 genome.

Bioinformatic notes and software used. In addition to the software noted above, Lasergene Seqbuilder and Seqman (DNASTAR) were used for sequencing and genome assembly. Mauve[27] was used extensively during the dynamic process of contig assembly and orientation. Analysis of RscS domain structure was assisted by PFAM[43] and Phobius[44]. Primer design was aided by Primer3Plus[33]. TreeView 1.6.6 (ref. 45) was used to view phylogenetic trees.

28. Stabb, E. V. & Ruby, E. G. RP4-based plasmids for conjugation between *Escherichia coli* and members of the Vibrionaceae. *Methods Enzymol.* **358,** 413–426 (2002).
29. Stabb, E. V. *et al.* in *Recent Advances in Marine Science and Technology 2000* (ed. Saxena, N.) 269–277 (PACON International, 2001).
30. Silhavy, T. J., Berman, M. L. & Enquist, L. W. *Experiments with Gene Fusions* (Cold Spring Harbor Laboratory Press, 1984).
31. Hussa, E. A., O'Shea, T. M., Darnell, C. L., Ruby, E. G. & Visick, K. L. Two-component response regulators of *Vibrio fischeri*: identification, mutagenesis, and characterization. *J. Bacteriol.* **189,** 5825–5838 (2007).
32. van Hijum, S. A., Zomer, A. L., Kuipers, O. P. & Kok, J. Projector 2: contig mapping for efficient gap-closure of prokaryotic genome sequence assemblies. *Nucleic Acids Res.* **33,** W560–W566 (2005).
33. Untergasser, A. *et al.* Primer3Plus, an enhanced web interface to Primer3. *Nucleic Acids Res.* **35,** W71–W74 (2007).
34. Altschul, S. F. *et al.* Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25,** 3389–3402 (1997).
35. Ruby, E. G. & Nealson, K. H. Pyruvate production and excretion by the luminous marine bacteria. *Appl. Environ. Microbiol.* **34,** 164–169 (1977).
36. Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F. & Higgins, D. G. The CLUSTAL X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* **25,** 4876–4882 (1997).
37. Tamura, K., Dudley, J., Nei, M. & Kumar, S. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol. Biol. Evol.* **24,** 1596–1599 (2007).
38. Swofford, D. L. *PAUP*: Phylogenetic Analysis Using Parsimony (* and Other Methods)* 4th edn (Sinauer, 2003).
39. Posada, D. & Crandall, K. A. MODELTEST: testing the model of DNA substitution. *Bioinformatics* **14,** 817–818 (1998).
40. Ronquist, F. & Huelsenbeck, J. P. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **19,** 1572–1574 (2003).
41. Zwickl, D. J. *Genetic Algorithm Approaches for the Phylogenetic Analysis of Large Biological Sequence Datasets under the Maximum Likelihood Criterion.* PhD. thesis, Univ. Texas at Austin (2006).
42. Karlin, S. & Mrázek, J. Predicted highly expressed genes of diverse prokaryotic genomes. *J. Bacteriol.* **182,** 5238–5250 (2000).
43. Finn, R. D. *et al.* Pfam: clans, web tools and services. *Nucleic Acids Res.* **34,** D247–D251 (2006).
44. Kall, L., Krogh, A. & Sonnhammer, E. L. Advantages of combined transmembrane topology and signal peptide prediction – the Phobius web server. *Nucleic Acids Res.* **35,** W429–W432 (2007).
45. Page, R. D. TreeView: an application to display phylogenetic trees on personal computers. *Comput. Appl. Biosci.* **12,** 357–358 (1996).

©2009 Macmillan Publishers Limited. All rights reserved

LETTERS

Cdc14 inhibits transcription by RNA polymerase I during anaphase

Andrés Clemente-Blanco[1], María Mayán-Santos[1], David A. Schneider[2], Félix Machín[1]†, Adam Jarmuz[1], Herbert Tschochner[3] & Luis Aragón[1]

Chromosome condensation and the global repression of gene transcription[1] are features of mitosis in most eukaryotes. The logic behind this phenomenon is that chromosome condensation prevents the activity of RNA polymerases. In budding yeast, however, transcription was proposed to be continuous during mitosis[2]. Here we show that Cdc14, a protein phosphatase required for nucleolar segregation[3] and mitotic exit[4], inhibits transcription of yeast ribosomal genes (rDNA) during anaphase. The phosphatase activity of Cdc14 is required for RNA polymerase I (Pol I) inhibition *in vitro* and *in vivo*. Moreover Cdc14-dependent inhibition involves nucleolar exclusion of Pol I subunits. We demonstrate that transcription inhibition is necessary for complete chromosome disjunction, because ribosomal RNA (rRNA) transcripts block condensin binding to rDNA, and show that bypassing the role of Cdc14 in nucleolar segregation requires *in vivo* degradation of nascent transcripts. Our results show that transcription interferes with chromosome condensation, not the reverse. We conclude that budding yeast, like most eukaryotes, inhibit Pol I transcription before segregation as a prerequisite for chromosome condensation and faithful genome separation.

In most eukaryotic cells, a dramatic structural reorganization of the genetic material into highly condensed chromosomes and the global silencing of gene transcription[1] accompany entry into mitosis. This might reflect an incompatibility between transcription and chromosome condensation and/or segregation processes. The most highly transcribed regions in eukaryotic genomes are the ribosomal gene arrays (rDNA), which require a dedicated polymerase named RNA Pol I. A study in budding yeast using cell size as an indicator of cell-cycle stage and cellular RNA transcript levels established that transcription, including rDNA, is not inhibited at any stage during mitosis[2]. This is surprising because rDNA becomes hyper-condensed during anaphase[5,6]. We thus revisited whether transcription is inhibited during yeast mitosis using more sensitive assays.

We measured total RNA synthesis in synchronised yeast cultures undergoing an entire cell cycle using incorporation of [^{3}H]uracil into total RNA and found that cells downregulate RNA synthesis during anaphase (Fig. 1a). Analysis of nascent 35S rRNA transcripts also showed significant reduction during anaphase (Fig. 1b; 75 min and Supplementary Fig. 1). The anaphase inhibition of rRNA transcription correlates with the exclusion of the Pol I subunit Rpa43 from the 35S gene region (Fig. 1d and Supplementary Figs 2 and 3). Therefore, yeast cells, like most eukaryotes, inhibit transcription during mitosis; however, whereas transcription inhibition in most eukaryotic cells takes place in metaphase, in yeast it occurs during anaphase.

In early anaphase the conserved phosphatase Cdc14 becomes activated[7]. Cdc14 is required for the resolution of transcription-dependent linkages in the ribosomal gene array[8,9]. Expression of *CDC14* from the *GAL1-10* promoter in metaphase cells causes a four-fold reduction in rRNA synthesis (Fig. 2a) and the effect is dependent on its phosphate activity (Fig. 2a). Cfi1 (also known as Net1), the nucleolar inhibitor of Cdc14 and a *CDK1* target[10–12], interacts directly with Pol I and stimulates transcription *in vitro*[13]. Expression of a stabilized form of cyclin B2, *CLB2-DK*[14], causes a mitotic arrest where Cdc14 is maintained at the fully released and active stage[12]. *CLB2-DK* expression also causes reduction in rRNA synthesis in the presence of wild-type Cdc14 (Fig. 2b), confirming that rDNA transcription inhibition is directly dependent on the phosphatase and not Cfi1. Moreover, *CLB2-DK* expression in the presence of Cdc14 causes delocalization of the essential Pol I subunit Rpa43 from the nucleolus (Fig. 2c, d and Supplementary Fig. 4). However, this is not the case for the entire Pol I holocomplex because the Rpa190 subunit is not delocalized when *CLB2-DK* is expressed in the presence of Cdc14 (Supplementary Fig. 5), despite the fact that Pol I transcription is inhibited (Fig. 2b). Therefore Cdc14 probably inhibits Pol I transcription by destabilization of specific subunits. These findings suggest that Cdc14 is a Pol I transcriptional repressor. Indeed, purified Cdc14 inhibits Pol I transcription *in vitro* (Fig. 3a) whereas the phosphatase-dead mutant does not (Fig. 3a). Cdc14 does not prevent stimulation Pol I transcription by Cfi1 (Fig. 3b). Therefore the activities of Cfi1 and Cdc14 in the activation and repression of Pol I transcription are independent.

Phosphorylation of Pol I has been linked to active transcription[15]. In metaphase-arrested cells at least two phospho-forms of Rpa43 can be detected (Supplementary Fig. 6). Expression of Cdc14 causes the loss of one of the Rpa43 phospho-bands (Supplementary Fig. 6). Therefore Cdc14 promotes dephosphorylation of Rpa43. Deletions in the carboxy-terminal region of Rpa43, where most phospho-sites are located[16], prevent Cdc14-dependent nucleolar delocalization but do not affect transcription inhibition (Supplementary Fig. 7). Therefore additional Cdc14 targets in the regulation of Pol I transcription repression must exist. Nonetheless, Rpa43 seems to be a key factor because a phospho-mimicking mutant of Rpa43 shows defects in transcription inhibition (Supplementary Fig. 8).

The requirement of Cdc14 for rDNA condensation[17,18] and segregation[17–20] is thought to stem from a defect in condensin localization[17,20]. Condensin binding is incompatible with rDNA transcription[21], and inactivation of Pol I allows rDNA segregation in the absence of Cdc14 (refs 8 and 9). Therefore transcription inhibition by Cdc14 is probably sufficient to promote condensin binding to rDNA. Indeed, in the absence of Cdc14 removal of rRNA transcripts by expression of the RntA RNase T1 from *Aspergillus oryzae*[22] (Supplementary Fig. 9) promotes condensin binding to rDNA (Fig. 4a).

[1]Cell Cycle Group, MRC Clinical Sciences Centre, Imperial College, Du Cane Road, London W12 0NN, UK. [2]Department of Biochemistry and Molecular Genetics, University of Alabama at Birmingham, 442 Kaul Human Genetics Building, 720 20th Street South, Birmingham, Alabama 35294, USA. [3]Institut für Biochemie, Mikrobiologie und Genetik, Universität Regensburg, Munich, Germany. †Present address: Unidad de Investigación, Hospital Universitario Candelaria, Carretera del Rosario, 38010 Tenerife, Spain.

©2009 Macmillan Publishers Limited. All rights reserved

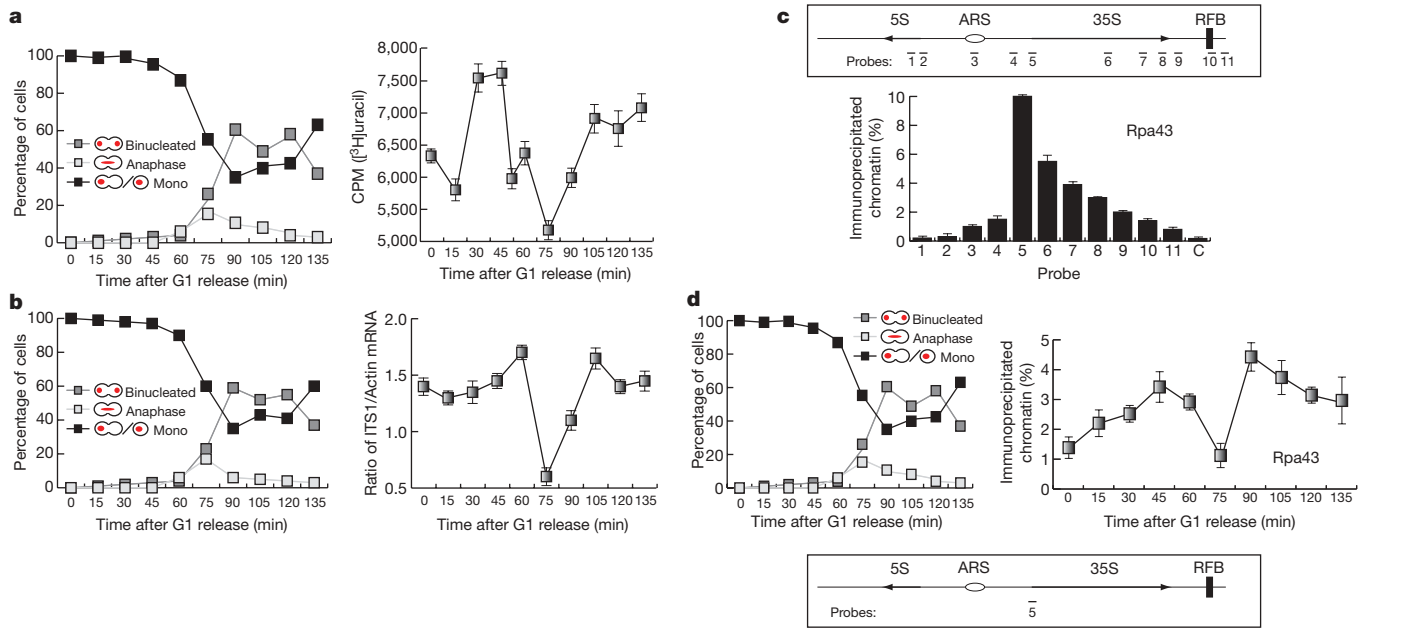


Figure 1 | Transcription is inhibited in budding yeast during anaphase. **a**, [^{3}H]uracil incorporation into total RNA in wild-type yeast cells released from a G1 block. The mean ($n = 3$) and s.d. are shown. **b**, Primary rRNA transcript levels in cells released from a G1 block. qPCR using primers to the internal transcribed sequence 1 (ITS1) (Supplementary Fig. 1) were used to determine levels of intact 35S rRNA transcript. An average ($n = 3$) and s.d. are shown. **c**, ChIP analysis of Rpa43–9myc binding to ribosomal repeats in exponentially growing cells. An average ($n = 2$) and s.d. are shown. **d**, ChIP analysis of Rpa43–9myc binding to the 5′ end region of the 35S rRNA gene in wild-type yeast cells released from a G1 block. An average ($n = 2$) and s.d. are shown.

Inactivation of cohesin in metaphase-arrested cells causes separation of all genomic regions except the rDNA because Cdc14 is inactive at this stage[17,18]. Expression of RntA in metaphase cells where cohesin has been inactivated (or a *cdc14-1* arrest; Supplementary Fig. 10) causes separation of nucleolar regions (Fig. 4b). Recent work has shown that expression of topoisomerase II from *Paramecium bursaria* Chlorella virus 1 (TopChlv), suppresses nucleolar segregation defects in condensin mutants[23]. In contrast, TopChlv does not promote separation of rDNA regions in metaphase cells inactivated for cohesin (Fig. 4b). This result is unexpected because Cdc14 and condensin were anticipated to be in the same pathway for rDNA segregation; thus the physical nature of rDNA linkages in the absence of either of these two factors should be identical. This is clearly not the case, as expression of RntA in metaphase cells where both cohesin and condensin are inactive does not rescue rDNA segregation (Fig. 4c).

An important feature of eukaryotic chromosomes is the pairing of replicated chromatids until metaphase and their separation in anaphase. In budding yeast, it was known that rDNA regions require

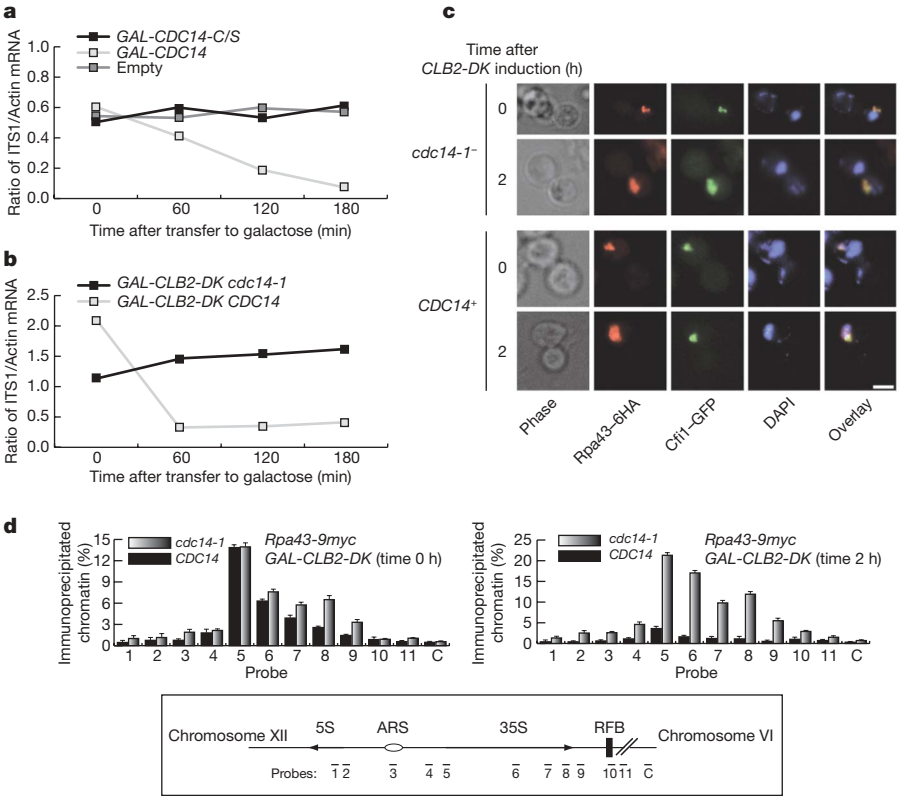


Figure 2 | Cdc14 phosphatase inhibits rRNA transcription and prevents binding of Pol I subunits to ribosomal genes. **a**, Primary rRNA transcript levels in metaphase-arrested cells with expression of wild-type *CDC14* or phosphatase-dead *CDC14-C/S*. **b**, Primary rRNA transcript levels upon expression of a stabilized form of cyclinB2 (*CLB2-DK*) in the presence of a wild-type *CDC14* or temperature-sensitive mutant *cdc14-1*. **c**, Localization of Rpa43 and the nucleolar marker Cfi1 in *CLB2-DK*-expressing cells with wild-type *CDC14* or *cdc14-1*. Scale bar 5 μm. **d**, ChIP analysis of Rpa43 binding to 35S DNA in cells before (left) and after (right) expression of *CLB2-DK* in the presence of a wild-type *CDC14* or *cdc14-1*. An average ($n = 2$) and s.d. are shown.

©2009 Macmillan Publishers Limited. All rights reserved

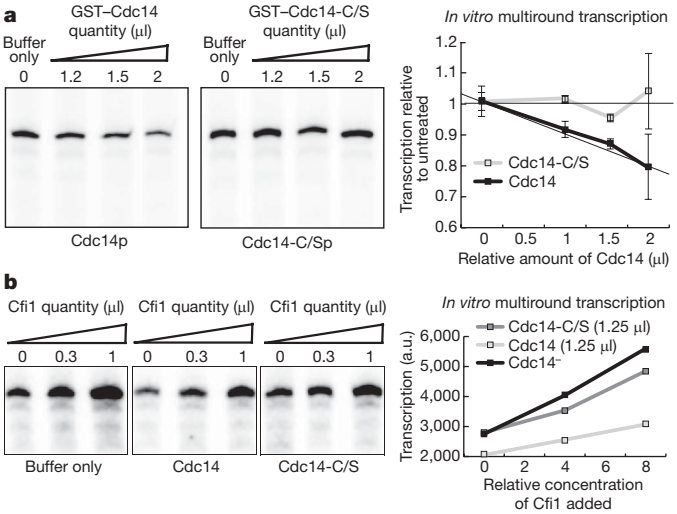


Figure 3 | Cdc14 inhibits RNA Pol I transcription *in vitro*. **a**, Multiround *in vitro* RNA Pol I transcription assays varying the relative amount of purified wild-type Cdc14 or phosphatase-dead Cdc14 (Cdc14-C/S) added to the reaction (concentration, 160 μg ml^{-1} of Cdc14 or Cdc14-C/S). Addition of Cdc14, but not Cdc14-C/S, inhibits RNA Pol I transcription. An average ($n = 2$) and s.d. are shown. **b**, Multiround *in vitro* RNA Pol I transcription assays varying the relative amount of Cfi1 added (concentration as in ref. 19) to reactions containing wild-type Cdc14, phosphatase-dead Cdc14-C/S or buffer only. a.u., arbitrary units.

Cdc14 (refs 17–20) and condensin[24] to segregate, in addition to cohesin cleveage[17,18]. However, the specific roles of these factors during segregation were not clear. Here, we have shown that Cdc14 function during rDNA segregation is the inhibition of transcription. Moreover, Cdc14 is a direct inhibitor of Pol I (Fig. 3a). We have shown that condensin localization to rDNA regions is dependent on this transcriptional inhibition, thus indirectly dependent on Cdc14. Condensin binding to rDNA is essential for compaction of the locus[5,24], a requirement for rDNA segregation[6]. Indeed, reducing the number of rDNA units partially bypasses the segregation requirement for Cdc14[8] (and hence downstream condensation). But are there additional requirements to transcription inhibition and condensation for rDNA segregation? Topological stress produced ahead of replication forks can be transferred to behind the forks by allowing replisome rotation, which generates precatenanes that become catenated daughter molecules after replication. In budding yeast, depletion of Top2, the enzyme that resolves catenations, causes severe segregation defects and 'cut phenotypes'[25]. In principle, DNA catenations within rDNA could also prevent its segregation. Indeed, some studies have suggested links between condensin and yeast Top2 (ref. 26); thus after transcription inhibition, condensin binding could mediate rDNA segregation by promoting not only condensation but also decatenation through Top2. Whether condensin function at rDNA includes Top2-mediated decatenation is an interesting question for the future.

Inhibition of transcription activity during mitosis was discovered many years ago[1]. The common explanation behind this phenomenon was that chromosome condensation limits the activity, although not

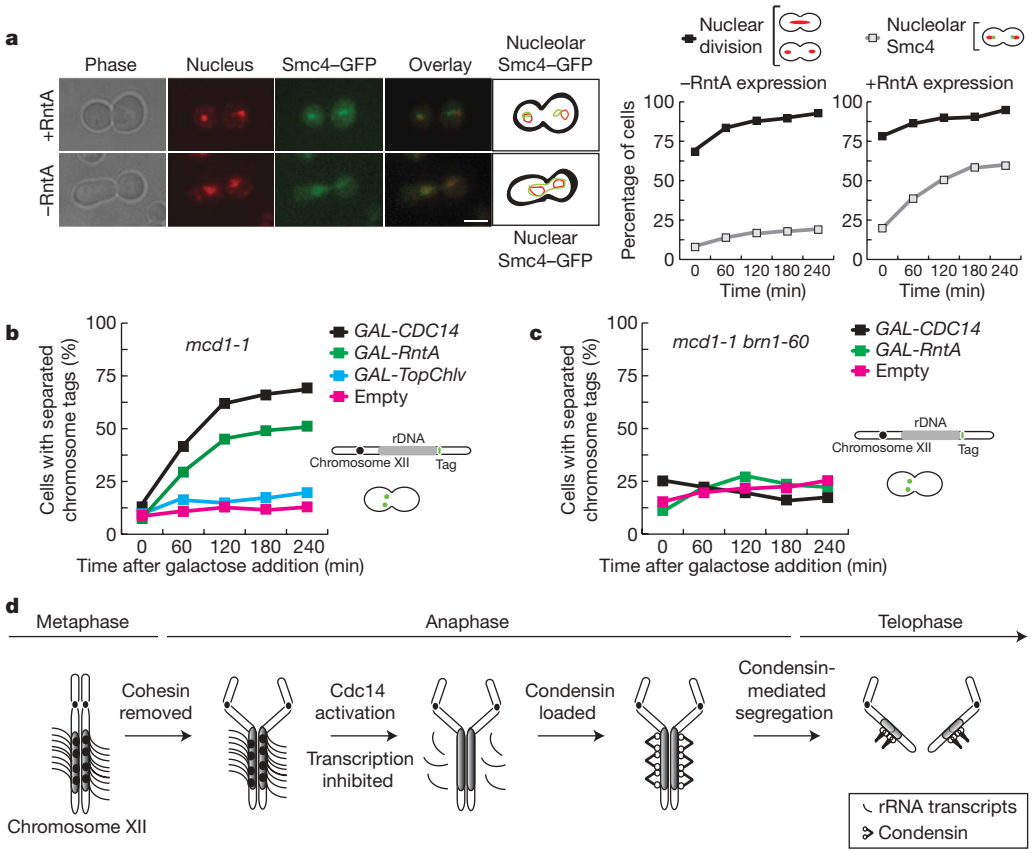


Figure 4 | Condensin localization to rDNA and nucleolar segregation requires removal of rRNA transcripts. **a**, Localization of the condensin subunit Smc4 in *cdc14-1* metaphase-arrested cells expressing the RntA RNase[22] (+RntA) or not (−RntA). Smc4 locates throughout the nucleoplasm (nuclear) or at the nucleolus (nucleolar). Representative micrographs and quantification of the experiment are shown. Scale bar, 5 μm. Two hundred cells were scored per time point. **b**, rDNA segregation in the cohesin mutant *mcd1-1* arrested in metaphase at 37 °C and expressing *CDC14*, RntA RNase, TopChlv or empty vector. Two hundred cells were scored per time point. **c**, rDNA segregation in the double cohesin condensin mutant *mcd1-1/brn1-60* arrested in metaphase at 37 °C and expressing *CDC14*, RntA or empty vector. Two hundred cells were scored per time point. **d**, Schematic representation of *CDC14*-dependent rRNA transcription inhibition during yeast anaphase.

©2009 Macmillan Publishers Limited. All rights reserved

the accessibility, of RNA polymerases to DNA. Our results show the opposite: the condensation complex (condensin) cannot access chromatin until rRNA transcription is significantly reduced. We show that in yeast rDNA this occurs during anaphase and is a process dependent on the mitotic exit phosphatase Cdc14 (Fig. 4d). Finally, we demonstrate that failure to reduce rRNA synthesis interferes with chromosome segregation. Therefore we conclude that mitotic inhibition of transcription is critical to maintain genome integrity through cell division.

METHODS SUMMARY

All media and growth conditions were as described[8]. The budding yeast strains used in this study are listed in Supplementary Methods. Growth, arrest and release conditions for every experiment shown in the manuscript figures are described in the Supplementary Methods. Details of gene tagging and constructions are described in Supplementary Methods. Microscope methods have been described[8]. ChIP was performed as described[27]. For multiround *in vitro* transcription assays, Pol I was prepared as described previously[28] from cells carrying His6-(HA)3 epitope tags on the amino terminus of A135 (NOY2173 and NOY2174). Multiround transcription with all pure components was performed as described at room temperature for 30 min[28]. Cdc14 and Cfi1 purification for *in vitro* assays were performed as described[13]. For quantitative PCR (qPCR) after reverse transcription (RT–PCR), RNAs were isolated with the RNeasy Mini Kit (Qiagen) using RNase-Free DNase Set (Qiagen) to digest DNA during RNA purification. Amplifications were done using Go Taq polymerase (Promega). Primer pairs and probes to the intergenic spacer region 1 (ITS1) were used to detect primary 35S rRNA (intergenic spacer sequences are removed from the primary 35S rRNA transcripts during the initial processing steps) (details in Supplementary Fig. 1). Quantifications were done using a Stratagene MX3000P Real-time PCR System and were confirmed by serial dilution analysis of RT–PCR reactions.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 30 June; accepted 13 November 2008.
Published online 21 January 2009.

1. Taylor, J. Nucleic acid synthesis in relation to the cell division cycle. *Ann. NY Acad. Sci.* **90,** 409–421 (1960).
2. Elliott, S. G. & McLaughlin, C. S. Regulation of RNA synthesis in yeast. III. Synthesis during the cell cycle. *Mol. Gen. Genet.* **169,** 237–243 (1979).
3. Granot, D. & Snyder, M. Segregation of the nucleolus during mitosis in budding and fission yeast. *Cell Motil. Cytoskeleton* **20,** 47–54 (1991).
4. Stegmeier, F. & Amon, A. Closing mitosis: the functions of the Cdc14 phosphatase and its regulation. *Annu. Rev. Genet.* **38,** 203–232 (2004).
5. Lavoie, B. D., Hogan, E. & Koshland, D. *In vivo* requirements for rDNA chromosome condensation reveal two cell-cycle-regulated pathways for mitotic chromosome folding. *Genes Dev.* **18,** 76–87 (2004).
6. Machin, F., Torres-Rosell, J., Jarmuz, A. & Aragon, L. Spindle-independent condensation-mediated segregation of yeast ribosomal DNA in late anaphase. *J. Cell Biol.* **168,** 209–219 (2005).
7. Stegmeier, F., Visintin, R. & Amon, A. Separase, polo kinase, the kinetochore protein Slk19, and Spo12 function in a network that controls Cdc14 localization during early anaphase. *Cell* **108,** 207–220 (2002).
8. Machin, F. *et al.* Transcription of ribosomal genes can cause nondisjunction. *J. Cell Biol.* **173,** 893–903 (2006).
9. Tomson, B. N., D'Amours, D., Adamson, B. S., Aragon, L. & Amon, A. Ribosomal DNA transcription-dependent processes interfere with chromosome segregation. *Mol. Cell. Biol.* **26,** 6239–6247 (2006).
10. Shou, W. *et al.* Exit from mitosis is triggered by Tem1-dependent release of the protein phosphatase Cdc14 from nucleolar RENT complex. *Cell* **97,** 233–244 (1999).
11. Visintin, R., Hwang, E. S. & Amon, A. Cfi1 prevents premature exit from mitosis by anchoring Cdc14 phosphatase in the nucleolus. *Nature* **398,** 818–823 (1999).
12. Azzam, R. *et al.* Phosphorylation by cyclin B-Cdk underlies release of mitotic exit activator Cdc14 from the nucleolus. *Science* **305,** 516–519 (2004).
13. Shou, W. *et al.* Net1 stimulates RNA polymerase I transcription and regulates nucleolar structure independently of controlling mitotic exit. *Mol. Cell* **8,** 45–55 (2001).
14. Hendrickson, C., Meyn, M. A. III, Morabito, L. & Holloway, S. L. The KEN box regulates Clb2 proteolysis in G1 and at the metaphase-to-anaphase transition. *Curr. Biol.* **11,** 1781–1787 (2001).
15. Fath, S. *et al.* Differential roles of phosphorylation in the formation of transcriptional active RNA polymerase I. *Proc. Natl Acad. Sci. USA* **98,** 14334–14339 (2001).
16. Gerber, J. *et al.* Site specific phosphorylation of yeast RNA polymerase I. *Nucleic Acids Res.* **36,** 793–802 (2008).
17. D'Amours, D., Stegmeier, F. & Amon, A. Cdc14 and condensin control the dissolution of cohesin-independent chromosome linkages at repeated DNA. *Cell* **117,** 455–469 (2004).
18. Sullivan, M., Higuchi, T., Katis, V. L. & Uhlmann, F. Cdc14 phosphatase induces rDNA condensation and resolves cohesin-independent cohesion during budding yeast anaphase. *Cell* **117,** 471–482 (2004).
19. Torres-Rosell, J., Machin, F., Jarmuz, A. & Aragon, L. Nucleolar segregation lags behind the rest of the genome and requires Cdc14p activation by the FEAR network. *Cell Cycle* **3,** 496–502 (2004).
20. Wang, B. D., Yong-Gonzalez, V. & Strunnikov, A. V. Cdc14p/FEAR pathway controls segregation of nucleolus in *S. cerevisiae* by facilitating condensin targeting to rDNA chromatin in anaphase. *Cell Cycle* **3,** 960–967 (2004).
21. Wang, B. D., Butylin, P. & Strunnikov, A. Condensin function in mitotic nucleolar segregation is regulated by rDNA transcription. *Cell Cycle* **5,** 2260–2267 (2006).
22. Fujii, T., Yamaoka, H., Gomi, K., Kitamoto, K. & Kumagai, C. Cloning and nucleotide sequence of the ribonuclease T1 gene (rntA) from *Aspergillus oryzae* and its expression in *Saccharomyces cerevisiae* and *Aspergillus oryzae*. *Biosci. Biotechnol. Biochem.* **59,** 1869–1874 (1995).
23. D'Ambrosio, C., Kelly, G., Shirahige, K. & Uhlmann, F. Condensin-dependent rDNA decatenation introduces a temporal pattern to chromosome segregation. *Curr. Biol.* **18,** 1084–1089 (2008).
24. Freeman, L., Aragon-Alcaide, L. & Strunnikov, A. The condensin complex governs chromosome condensation and mitotic transmission of rDNA. *J. Cell Biol.* **149,** 811–824 (2000).
25. Baxter, J. & Diffley, J. F. Topoisomerase II inactivation prevents the completion of DNA replication in budding yeast. *Mol. Cell* **30,** 790–802 (2008).
26. Bhalla, N., Biggins, S. & Murray, A. W. Mutation of YCS4, a budding yeast condensin subunit, affects mitotic and nonmitotic chromosome behavior. *Mol. Biol. Cell* **13,** 632–645 (2002).
27. Nelson, J. D., Denisenko, O. & Bomsztyk, K. Protocol for the fast chromatin immunoprecipitation (ChIP) method. *Nature Protocols* **1,** 179–185 (2006).
28. Keener, J., Josaitis, C. A., Dodd, J. A. & Nomura, M. Reconstitution of yeast RNA polymerase I transcription *in vitro* from purified components. TATA-binding protein is not required for basal transcription. *J. Biol. Chem.* **273,** 33795–33802 (1998).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We are grateful to A. Amon, D. Morgan, K. Kitamoto and J. Van Etten for providing reagents, and M. Nomura for advice. We thank members of the Aragon laboratory for discussions. We acknowledge grant support from UAB HSF-GEF to D.A.S. and from 'La Junta de Extremadura' to A.C.-B. The Aragon laboratory is supported by the Medical Research Council of the UK.

Author contributions F.M. made the initial observation that transcription impairs rDNA segregation in the absence of Cdc14 function. All experiments shown were performed by A.C.-B., except ChIPs which were performed by M.M.-S. and *in vitro* phosphatase assays which were performed by D.A.S. A.J. cloned the RNAse constructs. H.T. provided a battery of Rpa43 mutants. All experiments shown were conceived by L.A. and the person performing the specific experiment. L.A. wrote the paper.

Author information Correspondence and requests for materials should be addressed to L.A. (luis.aragon@csc.mrc.ac.uk).

©2009 Macmillan Publishers Limited. All rights reserved

METHODS

Yeast strains and plasmids. All yeast strains used are shown in Supplementary Table 1. Chromosome tags have been described elsewhere[6]. C-terminal epitope tagging of Cdc14, Smc4 and Rpa43 was performed using PCR allele replacement methods[29]. Plasmids for overexpression of Cdc14 and Top2 have been described previously[17]. *GAL-NLS-RntA-MYC* plasmid was created from pERT-1 (ref. 22; a gift from K. Kitamoto), adding an NLS sequence at the N terminus and a myc epitope at the C terminus (details available upon request). The glutathione *S*-transferase (GST)–Cdc14 and GST–Cdc14 phosphatase-dead mutant (*CDC14-C283S/R289*; *Cdc14-C/S*) were gifts from A. Amon and D. Morgan. *Top2* from *Chlorella virus* was amplified from genomic PBCV-1 (a gift from J. Van Etten). An NLS sequence at the N terminus was added to the *Top2* gene and a *GAL1-10* promoter for expression in yeast cells was cloned upstream of the ORF (details available upon request). Deletion of the C-terminal region of Rpa43 and point mutations were done using PCR allele replacement methods[29].

Cell-cycle experiments. Cells arrested in G1 were treated with α-factor as described previously[8]. To release cells from α-factor blocks, we transferred them to fresh media plus pronase E (0.1 mg ml^{-1}). For releases at non-permissive temperatures, we exposed cells to 37 °C for 30 min before their transfer to fresh media (also at 37 °C). Metaphase arrests using nocodazole were done incubating cells for 3 h with 15 μg ml^{-1} nocodazole. For the experiments of RntA overexpression in *cdc14* mutants, cells were inoculated overnight at 25 °C in –ura medium with 2% raffinose. Cells were transferred to medium containing 2% raffinose at 37 °C for 3 h to induce the arrest due to the inactivation of Cdc14. After microscopic confirmation of telophase arrest, the culture was incubated in medium containing 2% galactose at 37 °C to express RntA. For the experiments of RntA, *CDC14*, *CDC14-C/S* and *Top2* overexpression in *mcd1-1* or *mcd1-1 brn1-60* metaphase-arrested cells, cultures were arrested in metaphase (nocodazole) at 25 °C in 2% raffinose before shifting the temperature to 37 °C to inactivate cohesin (*mcd1-1*) and condensin (*brn1-60*) and adding 2% galactose to the media.

RNA methods. For RT–PCR analysis, RNAs were isolated with the RNeasy Mini Kit (Qiagen) using RNase-Free DNase Set (Qiagen) to digest DNA during RNA purification. One microgram of total RNA was used for the complementary DNA (cDNA) synthesis, using the Omniscript Reverse Transcription Kit (Qiagen). Finally, cDNA amplification was done using Go Taq polymerase (Promega). Primer pairs and probes to the intergenic spacer region 1 (ITS1) were used to detect primary 35S rRNA (intergenic spacer sequences are removed from the primary 35S rRNA transcripts during the initial processing steps). Quantifications were done using a Stratagene MX3000P Real-time PCR System and were confirmed by serial dilution analysis of RT–PCR reactions.

Microscopy and determination of segregation defects. Yeast cells with GFP-tagged proteins were analysed by fluorescence microscopy after 4,6-diamidino-2-phenylindole (DAPI) or propidium iodide staining. Images were collected on a Leica IRB using a Hamamatsu C4742-95 digital camera and OpenLab software (Improvision). Imaging was done in antifade medium (Molecular Probes) at room temperature. In each chromosome tag segregation/separation experiment, an average of 200 cells were counted for each time point and condition.

Cell extracts and western blotting. Collected cells were resuspended in 20 μl of RIPA buffer (10 mM sodium phosphate, 1% Triton X-100, 0.1% SDS, 10 mM EDTA, 150 mM NaCl, pH 7). Cells were disrupted in a Fast Prep machine (QBiogene). The total cell extract was recovered washing the beads with 200 μl of RIPA buffer. The soluble proteins were obtained by centrifugation for 5 min at 12,000 r.p.m. (4 °C). The protein concentration was measured using the BCA Protein Assay (Sigma). For SDS–PAGE gel electrophoresis, 30 μg of total cell extract was loaded per lane in 8% polyacrilamide gels (running in Tris-glicine buffer; Tris 3 g l^{-1}, glycine 14.4 g l^{-1}, SDS 1 g l^{-1}). Gels were transferred to Hybond-P membranes (Amersham-Bioscience) in Tris-glicine buffer without SDS. Membranes were blocked with 5% milk in PBS, 0.1% Tween-20 (Sigma). The Myc epitope was detected with an anti-Myc 9E10 monoclonal antibody (Roche). The haemagglutinin (HA) epitope was detected with anti-HA 12CA5 monoclonal antibody (Roche). For phosphatase assays, 30 μg of total cell extracts were incubated for 60 min at 30 °C with λ-phosphatase (New England Biolabs) before electrophoresis. For *in vitro* dephosphorylation assay, Rpa43-6HA was immunoprecipitated from cell extract containing 500 μg of total protein using μMACSTM Epitope Tag Protein Isolation Kit. Immunoprecipitates were incubated in phosphatase buffer (25 mM HEPES-NaOH; 150 mM NaCl; 0.1 mg ml^{-1} BSA; 2 mM $MnCl_2$) with GST–Cdc14, GST–Cdc14(C283S/R289A) (both purified from *Escherichia coli*) or λ-phosphatase. Proteins were analysed by immunoblotting with anti-HA antibodies.

Protein purification. *E. coli* cultures carrying the expression plasmids were pre-grown overnight (50 ml of lysogeny broth (LB) plus ampicillin) at 37 °C. One litre of LB medium was inoculated using the overnight culture and allowed to reach $OD_{600\,nm}$ of 1.0. Cells were induced with 0.5 M IPTG for 6 h. The pellet was re-suspended in 30 ml lysis buffer (0.1% Triton X-100, PBS and 1 mg ml^{-1} lysozyme) and incubated at room temperature for 15 min. Cells were sonicated three times for 30 s on ice, the extract was centrifuged at 14,000 r.p.m. for 30 min and the supernatant was collected. The supernatant was further incubated with 1 ml pre-equilibrated GST beads for 30 min at room temperature, centrifuged and washed three times with 5 ml lysis buffer, then twice with 5 ml lysis buffer plus 0.5 M NaCl and Triton X-100, followed by a final set of three washes with 5 ml lysis buffer. Elution was done by incubating with 5 ml of glutathione for 5 min. Purification of Cfi1 was done as in ref. 13.

Immunofluorescent staining. Cells were fixed with 3.7% formaldehyde for 1 h at room temperature. Cells were then washed and treated with 25 μg ml^{-1} zymolase 20T (in 0.1% β-mercaptoethanol and PBS) for 15 min at 37 °C. Next, cells were washed three times in PBS (sorbitol 1.2 M) and re-suspended in PBS (sorbitol 1.2 M, 1% Triton X-100) for 30 s. The Myc epitope was detected with an anti-Myc 9E10 monoclonal antibody (Roche) (diluted 1:500) overnight at 4 °C. Anti-mouse fluorescein-conjugated rabbit antibody (AbCam) was used for secondary detection (diluted 1:500).

Chromatin immunoprecipitation. ChIP analysis was done as in ref. 27. In brief, cells were fixed by treatment with 1% formaldehyde. Fixed cells were harvested, washed and re-suspended in ice-cold buffer I (140 mM NaCl, 1 mM EDTA, 1% (v/v) Triton X-100, 0.1% (w/v) sodium deoxycholate and 50 mM HEPES/KOH at pH 7.5) supplemented with protease inhibitor cocktail (Complete, Roche). Cells were broken with a FastPrep machine (Qbiogene) for four rounds of 30 s each with intervening incubations on ice for 5 min. Cell extracts were separated from the glass beads and sonicated (in a Bioruptor) for twelve rounds of 15 s with intervals of 45 s. This procedure fragments chromosomal DNA to an average size of about 500 base pairs. Purified anti-HA 12CA5 or anti-MYC 9e10 monoclonal antibodies (Roche) were used at 5 μg per 50 ml of initial culture. The antibodies were incubated with either protein A- or G-agarose beads (Roche) for several hours before use. Sonicated cell extracts were first pre-cleared with protein A- or G-agarose beads and then incubated overnight at 4 °C with the corresponding antibody beads. The beads were then washed three times for 15 min in buffer I, followed by three washes for 15 min in buffer II (500 mM NaCl, 1 mM EDTA, 1% (v/v) Triton X-100, 0.1% (w/v) sodium deoxycholate and 50 mM HEPES/KOH at pH 7.5) and three further washes also for 15 min in buffer III (250 mM LiCl, 1 mM EDTA, 0.5% (v/v) Nonidet P-40, 0.5% (w/v) sodium deoxycholate and 10 mM Tris-Cl at pH 8.0). All wash buffers were kept at 4 °C and supplemented with the above-described protease inhibitor cocktail. The beads were then transferred to a fresh tube and washed with TE (10 mM Tris at pH 7.6 and 1 mM EDTA). Immunoprecipitated material was incubated in elution buffer (1% SDS, 10 mM EDTA, 50 mM Tris-Cl at pH 8.0) at 65 °C for 10 min. The immunoprecipitated material was then separated from the beads and further incubated overnight at 65 °C to reverse the formaldehyde crosslinks. Aliquots of total chromatin solution were similarly heat-treated (whole-cell extract). After treatment with proteinase K (Qiagen), material was then treated with RNase (Qiagen) and finally purified through a DNA binding spin mini-column (Qiagen).

PCR analysis of co-immunoprecipitated DNA. The whole-cell extract (or total input) and immunoprecipitated samples were analysed by quantitative PCR as described above. Primer sequences are available on request. The values are given as a percentage of material immunoprecipitated (ChIP/input).

***In vitro* transcription.** Pol I was prepared as described previously[28] from cells carrying His6-(HA)3 epitope tags on the N terminus of A135 (NOY2173 and NOY2174). Multiround transcription with all pure components was performed as described at room temperature for 30 min (ref. 28).

***In vivo* labelling of RNA.** Labelling and analysis of RNA was performed as described[30]. In brief, cultures were grown in yeast peptone dextrose at 25 °C to $OD_{600\,nm} = 0.2$–0.3. Cultures were adjusted to $OD_{600\,nm} = 0.4$, pulse labelled with [3H]uracil (22 μCi ml^{-1} final; New England Nuclear Life Sciences/Perkin Elmer Life Sciences) for 5 min. The incorporation of [3H]uracil into RNA was measured using a scintillation counter.

29. Janke, C. *et al.* A versatile toolbox for PCR-based tagging of yeast genes: new fluorescent proteins, more markers and promoter substitution cassettes. *Yeast* **21**, 947–962 (2004).
30. Powers, T. & Walter, P. Regulation of ribosome biogenesis by the rapamycin-sensitive TOR-signaling pathway in *Saccharomyces cerevisiae*. *Mol. Biol. Cell* **10**, 987–1000 (1999).

©2009 Macmillan Publishers Limited. All rights reserved

LETTERS

Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals

Mitchell Guttman[1,2], Ido Amit[1], Manuel Garber[1], Courtney French[1], Michael F. Lin[1], David Feldser[3], Maite Huarte[1,6], Or Zuk[1], Bryce W. Carey[2,8], John P. Cassady[2,8], Moran N. Cabili[7], Rudolf Jaenisch[2,8], Tarjei S. Mikkelsen[1,4], Tyler Jacks[2,3], Nir Hacohen[1,9], Bradley E. Bernstein[1,10,11], Manolis Kellis[1,5], Aviv Regev[1,2], John L. Rinn[1,6,11]* & Eric S. Lander[1,2,7,8]*

There is growing recognition that mammalian cells produce many thousands of large intergenic transcripts[1–4]. However, the functional significance of these transcripts has been particularly controversial. Although there are some well-characterized examples, most (>95%) show little evidence of evolutionary conservation and have been suggested to represent transcriptional noise[5,6]. Here we report a new approach to identifying large non-coding RNAs using chromatin-state maps to discover discrete transcriptional units intervening known protein-coding loci. Our approach identified ~1,600 large multi-exonic RNAs across four mouse cell types. In sharp contrast to previous collections, these large intervening non-coding RNAs (lincRNAs) show strong purifying selection in their genomic loci, exonic sequences and promoter regions, with greater than 95% showing clear evolutionary conservation. We also developed a functional genomics approach that assigns putative functions to each lincRNA, demonstrating a diverse range of roles for lincRNAs in processes from embryonic stem cell pluripotency to cell proliferation. We obtained independent functional validation for the predictions for over 100 lincRNAs, using cell-based assays. In particular, we demonstrate that specific lincRNAs are transcriptionally regulated by key transcription factors in these processes such as p53, NFκB, Sox2, Oct4 (also known as Pou5f1) and Nanog. Together, these results define a unique collection of functional lincRNAs that are highly conserved and implicated in diverse biological processes.

There are at present only about a dozen well-characterized lincRNAs in mammals, with transcript sizes ranging from 2.3 to 17.2 kilobases (kb)[7,8]. These lincRNAs have distinctive biological roles through diverse molecular mechanisms, including functioning in X-chromosome inactivation (Xist, Tsix)[8,9], imprinting (H19, Air)[7,10], *trans*-acting gene regulation (HOTAIR)[11] and regulation of nuclear import (Nron)[12]. Importantly, these well-characterized lincRNAs show clear evolutionary conservation confirming that they are functional.

Genomic projects over the past decade have used shotgun sequencing and microarray hybridization[1–4] to obtain evidence for many thousands of additional non-coding transcripts in mammals. Although the number of transcripts has grown, so too have the doubts as to whether most are biologically functional[5,6,13]. The main concern was raised by the observation that most of the intergenic transcripts show little to no evolutionary conservation[5,13]. Strictly speaking, the absence of evolutionary conservation cannot prove the absence of function. But, the markedly low rate of conservation seen in the current catalogues of large non-coding transcripts (<5% of cases) is unprecedented and would require that each mammalian clade evolves its own distinct repertoire of non-coding transcripts. Instead, the data suggest that the current catalogues may consist largely of transcriptional noise, with a minority of bona fide functional lincRNAs hidden amid this background. Thus, to expand our understanding of functional lincRNAs, we are faced with two important challenges: (1) identifying lincRNAs that are most likely to be functional; and (2) inferring putative functions for these lincRNAs that can be tested in hypothesis-driven experiments.

To address the first challenge, we took an entirely different approach to discovering functional lincRNAs on the basis of exploiting chromatin structure. We recently developed an efficient method[14] to create genome-wide chromatin-state maps, using chromatin immunoprecipitation followed by massively parallel sequencing (ChIP-Seq). We observed that genes actively transcribed by RNA polymerase II (Pol II) are marked by trimethylation of lysine 4 of histone H3 (H3K4me3) at their promoter and trimethylation of lysine 36 of histone H3 (H3K36me3) along the length of the transcribed region[14]. We will refer to this distinctive structure as a 'K4–K36 domain'. We proposed that, by identifying K4–K36 structures that reside outside known protein-coding gene loci, we could systematically discover lincRNAs.

To test this hypothesis, we searched for K4–K36 domains in genome-wide chromatin-state maps of four mouse cell types: mouse embryonic stem cells (ESCs), mouse embryonic fibroblasts (MEF), mouse lung fibroblasts (MLF) and neural precursor cells (NPC). We identified K4–K36 domains of at least 5 kb in size that did not overlap regions containing protein-coding genes as well as known microRNAs[15] and endogenous short interfering RNAs (siRNAs)[16,17]. This analysis revealed 1,675 K4–K36 (1,250 conservatively defined) domains that do not overlap with known annotations; examples are shown in Fig. 1 (Supplementary Table 1).

Having identified K4–K36 loci with no previous annotation, we addressed: (1) whether these gene loci produce large multi-exonic RNA molecules; (2) whether the RNA molecules encode proteins or are non-coding transcripts; and (3) whether the RNA molecules, their promoters and their chromatin structure show conservation across mammals.

To test whether the intergenic K4–K36 domains produce RNA transcripts, we selected a random sample of 350 regions and designed

[1]Broad Institute of MIT and Harvard, 7 Cambridge Center, Cambridge, Massachusetts 02142, USA. [2]Department of Biology, [3]The Koch Institute for Integrative Cancer Research, [4]Division of Health Sciences and Technology, and [5]Computer Science and Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA. [6]Department of Pathology, Beth Israel Deaconess Medical Center, Boston, Massachusetts 02215, USA. [7]Department of Systems Biology, Harvard Medical School, Boston, Massachusetts 02114, USA. [8]Whitehead Institute for Biomedical Research, 9 Cambridge Center, Cambridge, Massachusetts 02142, USA. [9]Center for Immunology and Inflammatory Diseases, [10]Molecular Pathology Unit and Center for Cancer Research, Massachusetts General Hospital, Charlestown, Massachusetts 02129, USA. [11]Department of Pathology, Harvard Medical School, Boston, Massachusetts 02115, USA.
*These authors contributed equally to this work.

©2009 Macmillan Publishers Limited. All rights reserved

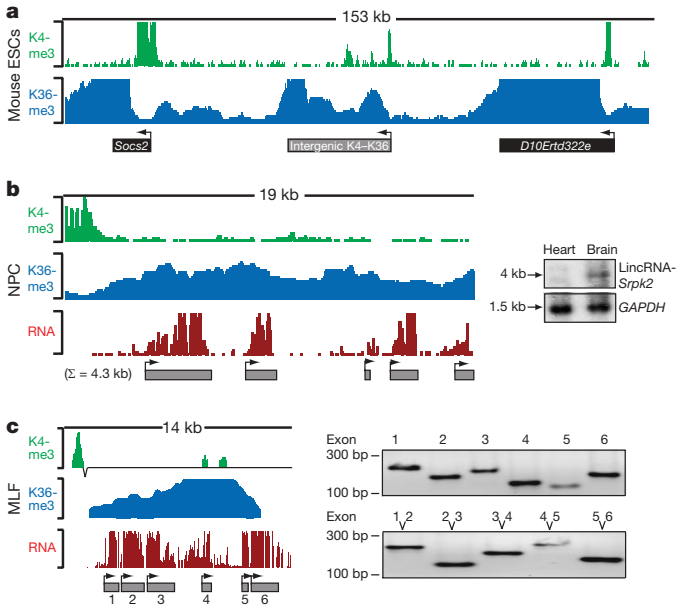


Figure 1 | Intergenic K4–K36 domains produce multi-exonic RNAs. **a**, Example of an intergenic K4–K36 domain and the K4–K36 domains of two flanking protein-coding genes. Each histone modification is plotted as the number of DNA fragments obtained by ChIP-Seq at each position. Black boxes indicate known protein-coding regions and grey boxes are intergenic K4–K36 domains. Arrowheads indicate the orientation of transcription. **b**, Intergenic K4–K36 domains were interrogated for presence of transcription by hybridizing RNA to DNA tiling arrays. The RNA hybridization intensity is plotted in red. RNA peaks were determined and are represented by grey boxes. The presence of a spliced transcript was validated by hybridization to a northern blot (right). **c**, Connectivity between the inferred exons was validated by PCR with reverse transcription (RT–PCR). Right top shows RT–PCR validation of each exon, right bottom shows RT–PCR across each consecutive exon.

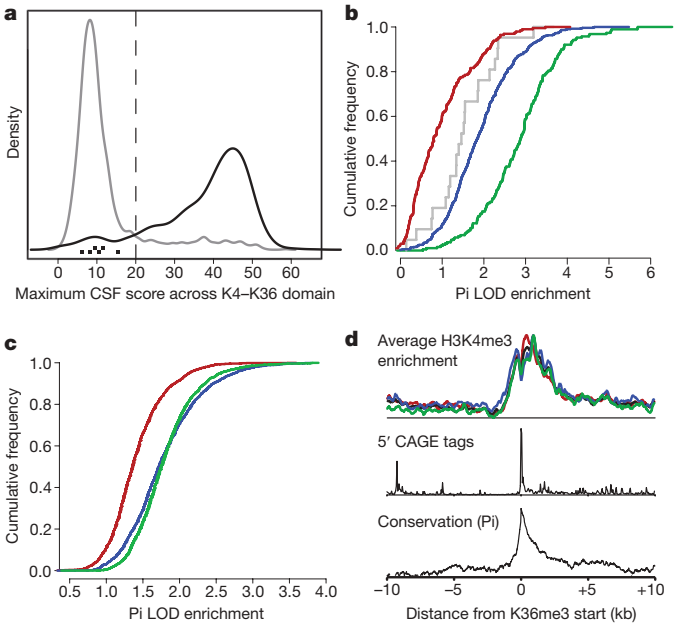


Figure 2 | lincRNA K4–K36 domains do not encode proteins and are conserved in their exons and promoters. **a**, Density plot of the maximum CSF score (Methods) across intergenic K4–K36 domains (grey) and known protein-coding genes (black). The maximum CSF scores for known lincRNAs are indicated as black points at the bottom. **b**, Cumulative distribution of sequence conservation across mammals for lincRNA exons (blue), protein-coding exons (green), introns (red) and known non-coding RNA exons (grey). **c**, Cumulative distribution of sequence conservation for lincRNA promoters (blue), random intergenic regions (red), and protein-coding promoters (green). LOD, logarithm of the odds ratio; Pi is the conservation metric (see Supplementary Methods). **d**, Enrichment of various promoter features plotted as the distance from the start of the K36me3 region averaged across all lincRNAs. Enrichment in each cell type of K4me3 domains across mouse ESCs (red), MEF (black), MLF (blue) and NPC (green) is shown (top panel). Enrichment of 5′ CAGE-tag density representing the 5′ end of RNA molecules (middle panel) and conservation scores in the K4me3 region are shown (bottom panel).

DNA microarrays containing oligonucleotides that tile across the regions (Methods) as well as various control regions. We hybridized poly(A)$^+$-selected RNA from each of the four cell types to the arrays. We developed an algorithm (Methods) to identify regions of significant hybridization and used it to define putative exons of transcripts detected at the loci. For ~70% of the intergenic loci with K4–K36 domains present in a cell type, we found clear evidence of RNA transcription in that cell type (Fig. 1 and Supplementary Tables 2 and 3). The proportion is similar to that for the protein-coding genes: ~72% of K4–K36 domains corresponding to known protein-coding genes show significant hybridization ($P < 0.05$, Wilcoxon test). In addition, we confirmed the presence of 93 out of 107 (87%) randomly selected exons, representing at least one exon from 19 out of 20 K4–K36 domains tested. We also confirmed the connectivity of consecutive exons in 52 out of 67 (78%) cases, including one from each of 16 K4–K36 domains tested (Fig. 1c and Supplementary Table 4). Furthermore, we validated the presence of discrete transcripts by hybridization to RNA northern blots in 15 of 17 tested loci (Fig. 1b, Supplementary Fig. 1, Supplementary Table 5 and Methods).

To determine whether the transcripts encode previously unknown protein-coding genes or non-coding RNAs, we used an established metric (the codon substitution frequency, CSF[18,19]) to assess characteristic evolutionary signatures of protein-coding domains. Analysing both the overall genomic locus (Fig. 2a and Supplementary Table 6) and the exons themselves (Supplementary Fig. 2 and Methods), we found that >90% of the intergenic K4–K36 domains fall well below the threshold of known protein-coding genes and resemble known lincRNAs (Fig. 2a). The result indicates that most of the loci do not encode protein-coding genes. Consistent with this, fewer than 2.5% of the exons show any similarity to known protein-coding genes, using the BLASTX program (Methods).

To assess the extent of nucleotide sequence conservation in the RNA transcripts, we used a method that explicitly models the underlying substitution rate (Methods) across 21 mammalian genomes (M.G. and X. Xie, submitted, Methods). We found that the lincRNA exons show clear sequence conservation when compared to other intergenic regions (Fig. 2b, Supplementary Fig. 3 and Supplementary Tables 2 and 7). Furthermore, the transcribed regions are highly enriched for conserved elements (defined by the PhastCons program[20]) compared to other intergenic regions ($P < 0.0001$, permutation test). The conservation level is similar to that seen for known lincRNAs, although it is lower than that seen for protein-coding exons, probably reflecting a lower degree of constraint on RNA structures than on amino-acid codons. The presence of strong purifying selection provides firm evidence that most K4–K36-defined lincRNAs must be biologically functional in mammals.

We used the same method to assess the conservation of the lincRNAs promoters (marked by the K4 domain). The lincRNA promoter regions show strong conservation, being essentially indistinguishable from known protein-coding genes (Fig. 2c and Supplementary Table 8). Furthermore, the lincRNA promoters show a notable enrichment of 'CAGE tags' (obtained by capturing the 7-methylguanosine cap at the 5′-end of Pol II transcripts) that mark transcriptional start sites[21] (Fig. 2d). Most of the lincRNA promoters regions (85%) contain a significant cluster of CAGE tags, with the density tightly localized around the promoter. In addition, the

©2009 Macmillan Publishers Limited. All rights reserved

lincRNA promoters show strong enrichment for binding of RNA Pol II in mouse ESCs ($P < 2 \times 10^{-16}$; Supplementary Fig. 4).

To investigate whether the K4–K36 chromatin structures observed at the loci are conserved across species, we constructed chromatin-state maps in human lung fibroblasts and MLF. Notably, ~70% of the K4–K36 domains in human also had a K4–K36 domain in the orthologous region of the mouse genome (Supplementary Table 9). The proportion is similar to that seen for protein-coding genes (~80%).

Together, the results show that most of the K4–K36 domains encode multi-exonic, non-protein-coding transcripts and the loci show clear conservation of nucleotide sequence and chromatin structure. Moreover, transcription and processing of these lincRNAs appears to be similar to that for protein-coding genes—including Pol II transcription, 5′-capping and poly-adenylation.

Having identified a large set of conserved lincRNAs, the next important challenge is to develop a method to infer putative functions that can be tested experimentally. To this end, we began by creating an RNA expression compendium of both lincRNAs and protein-coding genes across a wide range of tissues. We hybridized poly-adenylated RNA from 16 mouse samples to a custom lincRNA array. The samples included the original four cell types (mouse ESCs, NPC, MEF and MLF), a time course of embryonic development (whole embryo, hindlimb and forelimb at embryonic days 9.5, 10.5 and 13.5), and four normal adult tissues (brain, lung, ovary and testis) (Supplementary Fig. 5 and Supplementary Table 10).

The expression data contains a wealth of information about the lincRNAs. As an example, we searched for lincRNAs with an expression pattern opposite to the known lincRNA *HOTAIR*. Notably, we found that the most highly anti-correlated lincRNA in the genome lies in the HOXC cluster, in the same euchromatic domain as *HOTAIR*; we call this lincRNA *Frigidair* (Fig. 3c). This suggests that *Frigidair* may repress *HOTAIR* or perhaps activate genes in the HOXD cluster.

To take a more systematic approach, we also analysed RNA expression data for protein-coding genes from published sources[14,22] and generated further data for the embryonic development time course. We clustered the lincRNA and protein-coding genes into sets with correlated expression patterns (Supplementary Fig. 6a). We used Gene Set Enrichment Analysis (GSEA) to construct a matrix of the association of each lincRNA with each of ~1,700 functional gene sets (Fig. 3a and Supplementary Tables 10 and 11)[23]. We next performed biclustering on the gene set matrix to identify sets of lincRNAs that are associated with distinct sets of functional categories[24]. This analysis revealed numerous sets of lincRNAs associated with distinct and diverse biological processes (Fig. 3a). These include cell proliferation, RNA binding complexes, immune surveillance, ESC pluripotency, neuronal processes, morphogenesis, gametogenesis and muscle development (Fig. 3b and Supplementary Fig. 7).

To assess the validity of the inferred functional associations, we examined the gene sets associated with HOTAIR. *HOTAIR* showed negative association with *HOXD* genes (false discovery rate (FDR) < 0.018) and positive association with 'Chang Serum Response' (FDR < 0.001), a known predictor of breast cancer metastasis[25]. Both results are consistent with the known properties of *HOTAIR*, including a role in breast cancer metastasis[11,26].

We then sought to obtain independent experimental validation of the inferred biological functions for many of the lincRNAs. We focused on three large clusters of lincRNAs associated with the p53-mediated DNA damage response in MEF, NFκB signalling in dendritic cells, and ESC pluripotency, on the basis of their expression pattern across tissues.

We exposed $p53^{+/+}$ and $p53^{-/-}$ MEF to a DNA damaging agent and profiled the resulting expression changes on our lincRNA microarray (Methods)[27]. We found 39 lincRNAs that were significantly induced in $p53^{+/+}$ but not in $p53^{-/-}$ cells (Methods, Supplementary Fig. 8 and Supplementary Table 12). Approximately half of these lincRNAs resided in the cluster associated with p53-mediated DNA damage response, confirming the validity of the functional inference ($P < 10^{-7}$). Notably, we found that the promoters of these 39 lincRNAs were significant enriched for the p53 *cis*-regulatory binding element (versus all lincRNA promoters, $P < 0.01$, Wilcoxon test; Supplementary Fig. 8 and Supplementary Table 13). This suggests that p53 directly binds and regulates the expression of at least some of these lincRNA genes.

We stimulated $CD11C^{+}$ bone-marrow-derived dendritic cells with a specific agonist of the Toll-like receptor Tlr4, which signals through NFκB. We found that 20 lincRNAs showed marked upregulation after Tlr4-stimulation (Supplementary Table 14). Consistent with the inferences described earlier, 80% of these induced lincRNAs resided in the cluster associated with NFκB signalling. The greatest change in expression was observed in a lincRNA that is located ~51 kb upstream of the protein-coding gene *COX2* (also known as *Ptgs2*), a critical inflammation mediator that is directly induced by NFκB on Tlr4 stimulation; we refer to this as lincRNA-*COX2*. We found that lincRNA-*COX2* is induced ~1,000-fold over the course of 12 h after Tlr4 stimulation (Fig. 3d). In contrast, stimulation of Tlr3, which signals through IRF3, led to only weak induction of lincRNA-*COX2* (Fig. 3d).

Using published data from mouse ESCs, we identified 118 lincRNAs in which the promoter loci were bound by the core transcription factors Oct4 and Nanog[28] (Supplementary Table 15). Of those represented on our expression array 72% resided in the cluster associated with pluripotency, again supporting the validity of the functional inference. We noticed that one of these lincRNAs, which is only expressed in ESCs, is located ~100 kb from the *Sox2* locus, which encodes another key transcription factor associated with pluripotency (Fig. 3e). We cloned the promoter of this locus (which we will refer to as lincRNA-*Sox2*) upstream of a luciferase reporter gene and transfected the construct into mouse cells transiently expressing *Oct4*, *Sox2*, or both, as well as several controls. We found that *Sox2* and *Oct4* were each sufficient to drive expression of this lincRNA promoter, and the expression of both *Oct4* and *Sox2* together caused synergistic increases in expression (Fig. 3f). To our knowledge, this is the first experimental validation of a lincRNA promoter being directly regulated by key transcription factors such as Sox2 and Oct4.

The ultimate proof-of-function will be to demonstrate that RNA-interference-mediated knockout of each lincRNAs has the predicted phenotypic consequences. Towards this end, we examined a recently published short hairpin RNA screen of (presumed) protein-coding genes to identify genes that regulate cell proliferation rates in mouse ESCs[29]. The screen involved genes and some unidentified transcripts that had been identified as expressed in ESCs and showing rapid decrease in expression after retinoic acid treatment. Of the top ten hits in the screen, one corresponded to a gene of unknown function. We discovered that this gene corresponds to one of our lincRNAs (located ~181 kb from *Enc1*) contained in both the 'cell cycle and cell proliferation' cluster (FDR < 0.001) and the 'ESC' cluster (FDR < 0.001; Supplementary Fig. 9 and Supplementary Table 16). This provides functional confirmation that this lincRNA has a direct role in cell proliferation in ESCs, consistent with the analysis above.

Our results address the two key issues in the study of lincRNAs. We show that chromatin structure can identify sets of lincRNAs that show a high degree of evolutionary conservation, indicating that they are biologically functional. (We do not exclude the possibility that lincRNAs identified by shotgun sequencing that fail to show conservation are nonetheless functional, but other evidence will be required to establish this point.) We also provide a functional genomics pipeline for inferring putative roles for lincRNAs. The approach suggested functional roles for 150 lincRNAs that we studied on microarrays, and the independent experiments provided support for the predicted pathways for ~85 lincRNAs. The pipeline thus provides a useful guide for hypothesis-driven functional studies.

A fundamental issue will now be to determine the biological functions and the mechanisms by which lincRNAs act. One clue may come from the distribution of lincRNAs across the genome. We

©2009 Macmillan Publishers Limited. All rights reserved

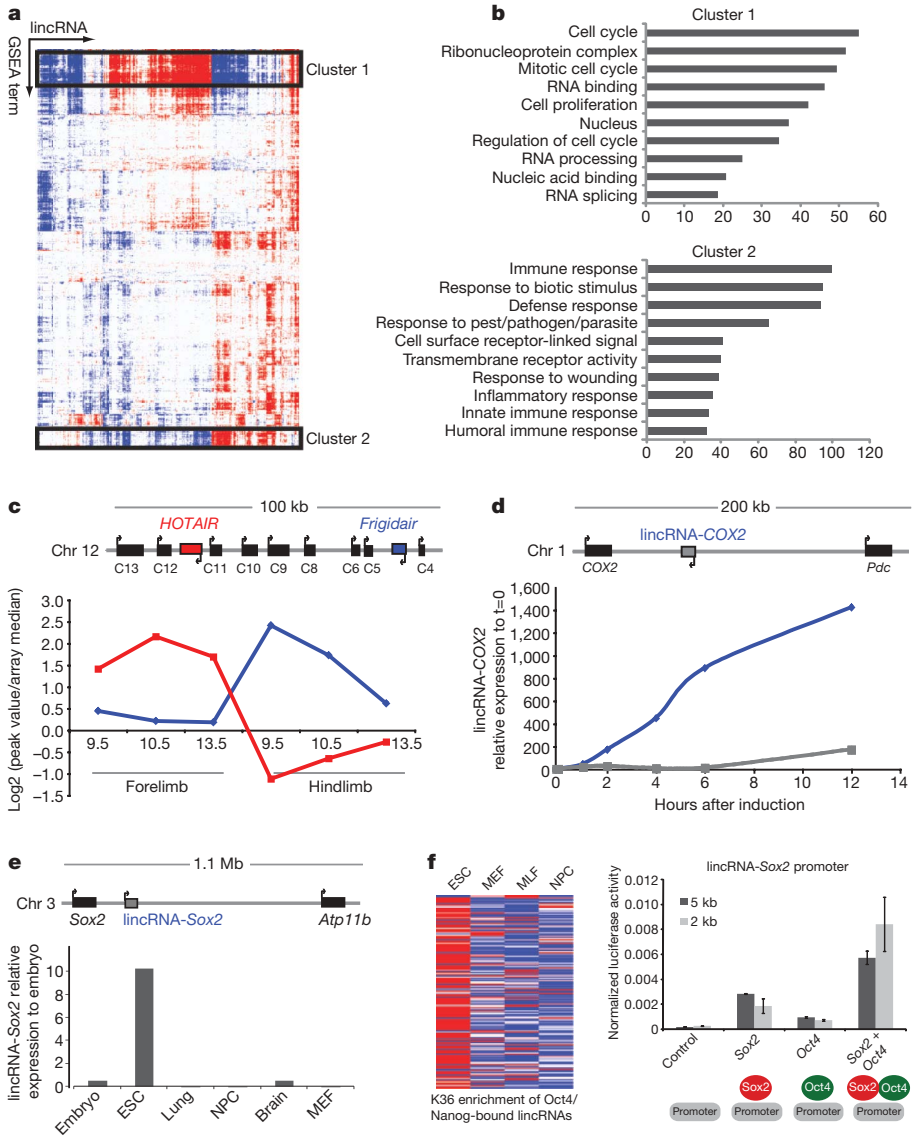


Figure 3 | lincRNAs show strong associations with other lincRNAs and with several biological processes. **a**, Association matrix of lincRNA and functional gene sets. Functional gene sets (columns) and lincRNAs (rows) are shown as positively (red), negatively (blue) or not associated (white) with lincRNA expression profiles. The black boxes highlight two significant biclusters in the matrix. **b**, Gene ontology of the protein-coding genes in these clusters is shown and plotted as the $-\log(P \text{ value})$ for the enrichment of each Gene Ontology term. **c**, Map of mouse genomic locus (*Hoxc*) containing *HOTAIR*. *HOTAIR* (red) and *Frigidair* (blue) show diametrically opposed expression patterns between mouse forelimb (anterior) and mouse hindlimb (posterior). **d**, Map of genomic locus containing *COX2* along with the location of lincRNA-*COX2*. Quantitative RT–PCR shows that lincRNA-*COX2* is upregulated in TLR4-stimulated cells (blue) but not TLR3-stimulated cells (grey). **e**, A map of the genomic locus containing *Sox2* shows a lincRNA ~50 kb upstream that is expressed specifically in ESCs. **f**, K36me3 enrichment across four cell types for lincRNAs bound by Oct4 or Nanog (left). Red indicates high enrichment, blue denotes low enrichment. The lincRNA-*Sox2* promoter was cloned into a luciferase reporter construct and assayed for transcriptional activity with *Sox2* and *Oct4* alone, together and controls (right). The *y*-axis represents the transcriptional activity of this promoter relative to a Renilla construct. Error bars are ± s.d. of three replicate transfections.

noted that several of the lincRNAs were located near genes encoding transcription factors (such as Sox2, Klf4, Myc and Brn1). Analysing the set of lincRNAs, we found that the genes neighbouring lincRNAs were strongly biased towards those encoding transcription factors ($P < 0.001$, permutation test; Supplementary Fig. 10 and Supplementary Table 17) and other proteins factors related to transcription. A second clue may come from our previous observation that HOTAIR[11] represses gene expression and is associated with chromatin remodelling proteins, together with recent similar observations for XIST[30]. On the basis of these observations, we speculate that many lincRNAs may be involved in transcriptional control—perhaps by guiding chromatin remodelling proteins to target loci—and that some transcription factors and lincRNAs may act together, with the transcription factor activating a transcriptional program and the lincRNA repressing a previous transcriptional program. Testing these speculations will require biochemical and genetic studies, including gene knockdown in appropriate settings. Whatever their functions, the highly conserved lincRNAs represent an important new contingent in the growing population of the modern 'RNA world'.

METHODS SUMMARY

Identifying intergenic K4–K36 domains and RNA. Enriched K4–K36 domains were identified using a sliding window approach across the genome and assessing the significance of each window. We filtered the list of K4–K36 enriched domains to eliminate known annotations. DNA tiling arrays (Nimblegen) were designed to tile intergenic K4–K36 domains. Transcribed regions were defined using a sliding window approach.

Conservation and coding potential. To detect sequence constraint we used a method that explicitly modelled the rate of mutation and level of constraint. We

©2009 Macmillan Publishers Limited. All rights reserved

took the maximum 12-base-pair window score for each exonic region. We normalized for the size differences between exons by computing a size matched random genomic score (Supplementary Methods).

We tested the protein-coding potential of K4–K36 domains by determining the maximum CSF score observed across the entire genomic locus. We computed the CSF scores across sliding windows of 90 base pairs and scanned all six possible reading frames in each window.

Protein-coding gene expression profiles. We generated a correlation matrix between lincRNAs and between lincRNAs and protein-coding genes by computing the Pearson correlation for all pairwise combinations. This matrix was clustered and visualized using the Gene Pattern platform for integrative genomics (http://genepattern.broad.mit.edu/). Functional associations were computed using Gene Set Enrichment Analysis (GSEA) (Supplementary Methods). In brief, we used each lincRNA as a profile, computed the Pearson correlation for each protein-coding gene and then ranked the protein-coding genes by their correlation coefficient. Gene sets were filtered by an FDR < 0.05 and an association matrix was generated.

Received 5 August; accepted 25 November 2008.
Published online 1 February 2009.

1. Bertone, P. *et al.* Global identification of human transcribed sequences with genome tiling arrays. *Science* **306,** 2242–2246 (2004).
2. Carninci, P. *et al.* The transcriptional landscape of the mammalian genome. *Science* **309,** 1559–1563 (2005).
3. Kapranov, P. *et al.* Large-scale transcriptional activity in chromosomes 21 and 22. *Science* **296,** 916–919 (2002).
4. Rinn, J. L. *et al.* The transcriptional activity of human chromosome 22. *Genes Dev.* **17,** 529–540 (2003).
5. Ponjavic, J., Ponting, C. P. & Lunter, G. Functionality or transcriptional noise? Evidence for selection within long noncoding RNAs. *Genome Res.* **17,** 556–565 (2007).
6. Struhl, K. Transcriptional noise and the fidelity of initiation by RNA polymerase II. *Nature Struct. Mol. Biol.* **14,** 103–105 (2007).
7. Brannan, C. I., Dees, E. C., Ingram, R. S. & Tilghman, S. M. The product of the *H19* gene may function as an RNA. *Mol. Cell. Biol.* **10,** 28–36 (1990).
8. Brown, C. J. *et al.* A gene from the region of the human X inactivation centre is expressed exclusively from the inactive X chromosome. *Nature* **349,** 38–44 (1991).
9. Lee, J. T., Davidow, L. S. & Warshawsky, D. *Tsix*, a gene antisense to *Xist* at the X-inactivation centre. *Nature Genet.* **21,** 400–404 (1999).
10. Sotomaru, Y. *et al.* Unregulated expression of the imprinted genes *H19* and *Igf2r* in mouse uniparental fetuses. *J. Biol. Chem.* **277,** 12474–12478 (2002).
11. Rinn, J. L. *et al.* Functional demarcation of active and silent chromatin domains in human *HOX* loci by noncoding RNAs. *Cell* **129,** 1311–1323 (2007).
12. Willingham, A. T. *et al.* A strategy for probing the function of noncoding RNAs finds a repressor of NFAT. *Science* **309,** 1570–1573 (2005).
13. Wang, J. *et al.* Mouse transcriptome: neutral evolution of 'non-coding' complementary DNAs. *Nature* **431,** 1–2, doi: 10.1038/nature03016 (2004).
14. Mikkelsen, T. S. *et al.* Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* **448,** 553–560 (2007).
15. Griffiths-Jones, S., Grocock, R. J., van Dongen, S., Bateman, A. & Enright, A. J. miRBase: microRNA sequences, targets and gene nomenclature. *Nucleic Acids Res.* **34,** D140–D144 (2006).
16. Tam, O. H. *et al.* Pseudogene-derived small interfering RNAs regulate gene expression in mouse oocytes. *Nature* **453,** 534–538 (2008).
17. Watanabe, T. *et al.* Endogenous siRNAs from naturally formed dsRNAs regulate transcripts in mouse oocytes. *Nature* **453,** 539–543 (2008).
18. Clamp, M. *et al.* Distinguishing protein-coding and noncoding genes in the human genome. *Proc. Natl Acad. Sci. USA* **104,** 19428–19433 (2007).
19. Lin, M. F. *et al.* Revisiting the protein-coding gene catalog of *Drosophila melanogaster* using 12 fly genomes. *Genome Res.* **17,** 1823–1836 (2007).
20. Siepel, A. *et al.* Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. *Genome Res.* **15,** 1034–1050 (2005).
21. Carninci, P. *et al.* Genome-wide analysis of mammalian promoter architecture and evolution. *Nature Genet.* **38,** 626–635 (2006).
22. Su, A. I. *et al.* Large-scale analysis of the human and mouse transcriptomes. *Proc. Natl Acad. Sci. USA* **99,** 4465–4470 (2002).
23. Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl Acad. Sci. USA* **102,** 15545–15550 (2005).
24. Tanay, A., Sharan, R. & Shamir, R. Discovering statistically significant biclusters in gene expression data. *Bioinformatics* **18** (Suppl 1), S136–S144 (2002).
25. Chang, H. Y. *et al.* Robustness, scalability, and integration of a wound-response gene expression signature in predicting breast cancer survival. *Proc. Natl Acad. Sci. USA* **102,** 3738–3743 (2005).
26. Carrio, M., Arderiu, G., Myers, C. & Boudreau, N. J. Homeobox D10 induces phenotypic reversion of breast tumor cells in a three-dimensional culture model. *Cancer Res.* **65,** 7177–7185 (2005).
27. Ventura, A. *et al.* Cre-lox-regulated conditional RNA interference from transgenes. *Proc. Natl Acad. Sci. USA* **101,** 10380–10385 (2004).
28. Loh, Y. H. *et al.* The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells. *Nature Genet.* **38,** 431–440 (2006).
29. Ivanova, N. *et al.* Dissecting self-renewal in stem cells with RNA interference. *Nature* **442,** 533–538 (2006).
30. Zhao, J., Sun, B. K., Erwin, J. A., Song, J. J. & Lee, J. T. Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome. *Science* **322,** 750–756 (2008).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We would like to thank our colleagues at the Broad Institute, especially J. P. Mesirov for discussions and statistical insights, X. Xie for statistical help with conservation analyses, J. Robinson for visualization help, M. Ku, E. Mendenhall and X. Zhang for help generating ChIP samples, and N. Novershtern and A. Levy for providing transcription factor lists. M. Guttman is a Vertex scholar, I.A. acknowledges the support of the Human Frontier Science Program Organization. This work was funded by Beth Israel Deaconess Medical Center, National Human Genome Research Institute, and the Broad Institute of MIT and Harvard.

Author Contributions J.L.R., E.S.L., A.R. and M. Guttman conceived and designed experiments. The manuscript was written by M. Guttman, A.R., J.L.R. and E.S.L. J.L.R., I.A., C.F., D.F., M.H., B.W.C., J.P.C. and M. Guttman performed molecular biology experiments. All data analyses were performed by M. Guttman in conjunction with M. Garber (conservation analyses), M.F.L. (codon substitution frequency), T.S.M. (ChIP-seq data), O.Z. (motif analysis) and M.N.C. (lincRNA genomic location analysis). Reagents were provided by M. Garber (pre-published conservation analysis tools); T.J. and D.F. (p53 wild-type and knockout MEFs); N.H., A.R. and I.A. (dendritic cell stimulated time course); B.E.B. (ChIP data); R.J., B.W.C. and J.P.C. (luciferase assays); and M.K. and M.F.L. (codon substitution frequency code).

Author Information Microarray data have been deposited in the Gene Expression Omnibus (GEO) under accession number GSE13765. Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to J.L.R. (jrinn@broad.mit.edu).

©2009 Macmillan Publishers Limited. All rights reserved

LETTERS

Ubiquitin-related modifier Urm1 acts as a sulphur carrier in thiolation of eukaryotic transfer RNA

Sebastian Leidel[1]*, Patrick G. A. Pedrioli[1]*, Tamara Bucher[1], Renée Brost[2], Michael Costanzo[2], Alexander Schmidt[3,6], Ruedi Aebersold[3,4,5,6], Charles Boone[2], Kay Hofmann[7] & Matthias Peter[1,6]

Ubiquitin-like proteins (UBLs) can change protein function, localization or turnover by covalent attachment to lysine residues[1]. Although UBLs achieve this conjugation through an intricate enzymatic cascade, their bacterial counterparts MoaD and ThiS function as sulphur carrier proteins[2,3]. Here we show that Urm1p, the most ancient UBL[3], acts as a sulphur carrier in the process of eukaryotic transfer RNA (tRNA) modification, providing a possible evolutionary link between UBL and sulphur transfer. Moreover, we identify Uba4p, Ncs2p, Ncs6p and Yor251cp as components of this conserved pathway. Using *in vitro* assays, we show that Ncs6p binds to tRNA, whereas Uba4p first adenylates and then directly transfers sulphur onto Urm1p. Finally, functional analysis reveals that the thiolation function of Urm1p is critical to regulate cellular responses to nutrient starvation and oxidative stress conditions, most likely by increasing translation fidelity.

Urm1p, an evolutionarily conserved UBL from *Saccharomyces cerevisiae*[2], exhibits sequence and structural similarities with the bacterial MoaD and ThiS-systems[3], but is thought to be conjugated to target proteins[4–6]. Initial characterization has identified Uba4p as its E1-enzyme and placed the *URM1* pathway in the functional context of invasive growth, response to oxidative stress, nutrient starvation and Tor-signalling[4–7]. Although the Urm1p target Ahp1p plays an essential role during oxidative stress[8], many phenotypes observed in *urm1*Δ cells cannot be explained by Ahp1p regulation, implying that additional Urm1p targets must exist.

To understand Urm1p molecular function better, we performed synthetic genetic array (SGA) analysis with *urm1*Δ and *uba4*Δ query-strains. Twenty-four genes were present in both screens and statistically enriched for the Gene Ontology annotation 'nucleic acid metabolism' (Fig. 1a and Supplementary Table 1). In particular, we identified all non-essential components of the elongator (ELP) complex, which is required for methoxy-carbonyl-methyl modification of wobble uridine$_{34}$ (mcm^5U$_{34}$) of certain tRNAs (Fig. 1b)[9]. Moreover, ELP complex[10], and *URM1* or *UBA4*[11], mutations confer resistance to the modification-dependent endonuclease γ-toxin from *Kluyveromyces lactis*[12], suggesting a functional connection between the ELP complex and the *URM1* pathway. Intriguingly, the U$_{34}$ residue in the ELP-targeted tRNAs tK^{UUU}, tE^{UUC} and tQ^{UUG} is further modified by a thiocarbonyl group (mcm^5s^2U) (Fig. 1b). Both modifications are important for translational fidelity and reading-frame maintenance[13]. As the pathway responsible for this second modification in the cytosol of eukaryotes is unknown, we speculated that Urm1p and Uba4p might be accountable for it.

Consistent with our hypothesis and results obtained for the ELP complex[12,14], tRNA overexpression rescued known growth defects of *urm1*Δ and *uba4*Δ strains[5,6,11] including sensitivity to the Tor inhibitors rapamycin and caffeine, or diamide, which reversibly oxidizes glutathione (Fig. 1c and Supplementary Fig. 1), providing strong genetic evidence for a role of tRNA in mediating these phenotypes. Interestingly, distinct tRNAs unequally rescued different phenotypes, most likely reflecting differences in codon usage of genes essential under the individual conditions.

((*N*-acryloylamino)phenyl)mercuric chloride (APM), which specifically retards migration of thio-containing molecules like mcm^5s^2U$_{34}$ through polyacrylamide gel electrophoresis (PAGE) gels[15], was used to compare tRNA thiolation in wild-type, *urm1*Δ and *uba4*Δ strains. Interestingly, a fraction of total tRNA isolated from wild-type but not *urm1*Δ and *uba4*Δ cells was shifted (Fig. 2a, b and Supplementary Fig. 2a), suggesting that the *URM1* pathway is required for tRNA thiolation. Northern blot analysis confirmed the specificity of this assay for mcm^5s^2U-containing tRNA species (Fig. 2b and Supplementary Fig. 2a).

To confirm unequivocally that the *URM1* pathway is essential for thiolation, we column-purified single tRNA species from wild-type, *urm1*Δ and *uba4*Δ cells and analysed the nucleosides by mass spectrometry (Supplementary Fig. 3). An elution peak corresponding to mcm^5s^2U was exclusively present in the wild-type samples (Fig. 2c and Supplementary Fig. 4a). As expected, *urm1*Δ and *uba4*Δ, but not *elp1*Δ samples, contained mcm^5U, suggesting that only thiolation was disrupted (Supplementary Figs 4b and 5). Furthermore, *elp1*Δ samples showed a significantly lower intensity of the s^2U-containing species by APM gel analysis (Fig. 2a, b and Supplementary Fig. 2a). Taken together, these results demonstrate that the *URM1* pathway is required for the s^2U modification of specific tRNAs and suggest that this thiolation is facilitated by previous mcm^5U formation. Eukaryotic cells contain separate tRNA pools in the cytosol and mitochondria (mt-tRNA). Interestingly, thiolation of mt-tK^{UUU}, mt-tE^{UUC} and mt-tQ^{UUG} was not altered in *urm1*Δ and *uba4*Δ cells, but was absent in *slm3*Δ cells, as previously reported[16] (Supplementary Fig. 2b, f). Conversely, cytoplasmic tK^{UUU} was thiolated in *slm3*Δ cells, indicating that *URM1* is specifically required to thiolate cytoplasmic tRNA.

We used genetic and proteomic approaches to identify additional *URM1* pathway components required for tRNA thiolation. First, we performed reciprocal SGA screens using *elp4*Δ and *elp6*Δ query-strains. Second, we purified His$_{10}$HA-Urm1p from yeast and subsequently analysed reciprocal tandem affinity purifications (TAPs) of putatively interacting proteins by mass spectrometry. Finally, we

[1]Institute of Biochemistry, ETH Zurich, Schafmattstrasse 18, CH-8093 Zurich, Switzerland. [2]Banting and Best Department of Medical Research and Department of Molecular Genetics, University of Toronto, Toronto, Ontario M5S 3E1, Canada. [3]Institute of Molecular Systems Biology, ETH Zurich, Wolfgang-Pauli-Strasse 16, CH-8093 Zurich, Switzerland. [4]Faculty of Science, University of Zurich, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland. [5]Institute of Systems Biology, 1441 North 34th Street, Seattle, Washington 98103, USA. [6]Competence Center for Systems Physiology and Metabolic Diseases, Schafmattstrasse 18, CH-8093 Zurich, Switzerland. [7]Bioinformatics Group, Miltenyi Biotec GmbH, Friedrich-Ebert-Strasse 68, 51429 Bergisch-Gladbach, Germany.
*These authors contributed equally to this work.

©2009 Macmillan Publishers Limited. All rights reserved

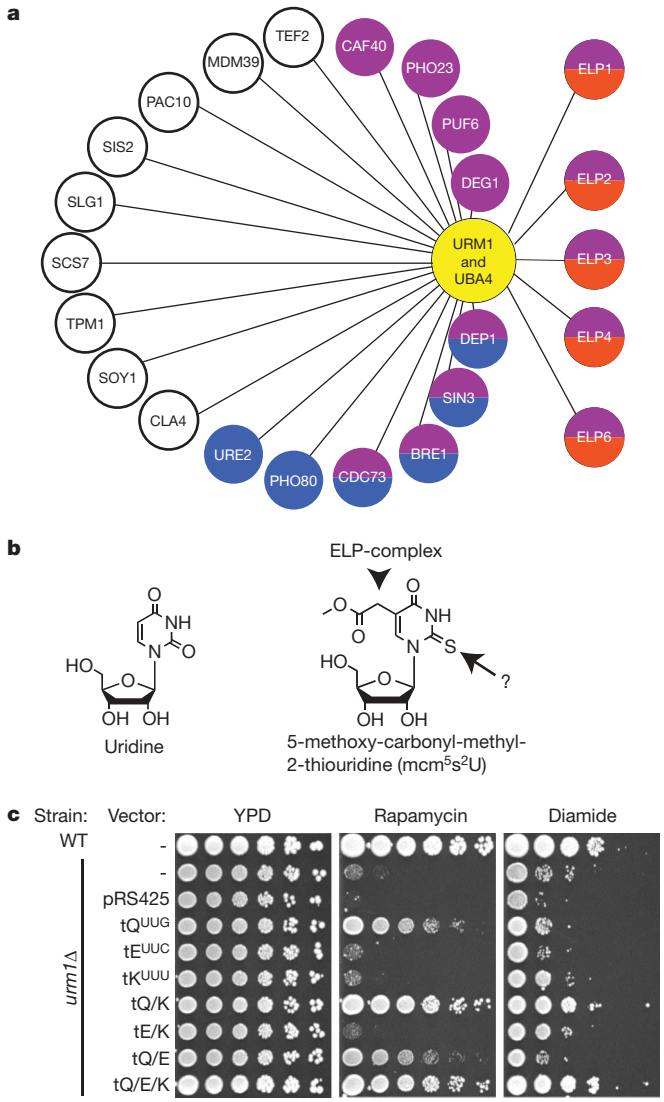


Figure 1 | Urm1p and Uba4p are functionally linked to the elongator complex. **a**, Genes identified by SGA screens with the query-strains *urm1*Δ and *uba4*Δ. Statistically overrepresented Gene Ontology categories are shown: tRNA modification (red), telomere organization and biogenesis (blue), and nucleobase, nucleoside, nucleotide and nucleic acid metabolism (violet). **b**, Chemical structure of uridine (top) and 5-methoxy-carbonyl-methyl-2-thiouridine (mcm^5s^2U) (bottom). The thiocarbonyl group is marked by an arrow, the methoxy-carbonyl-methyl by an arrowhead. **c**, Spot dilution assays of *urm1*Δ yeast expressing elevated levels of tE^{UUC}, tK^{UUU} and tQ^{UUG} either as single tRNAs or in combinations on YPD plates containing 2 nM rapamycin or 1 mM diamide.

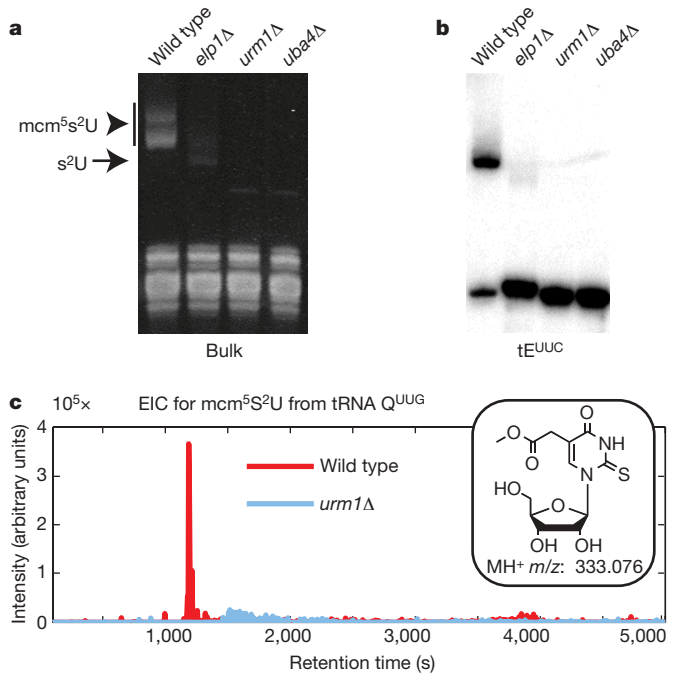


Figure 2 | Cytoplasmic tRNAs from *urm1*Δ and *uba4*Δ cells are not thiolated. **a**, **b**, ((*N*-acryloylamino)phenyl)mercuric chloride (APM)-gel retardation analysis of tRNA isolated from wild-type, *elp1*Δ, *urm1*Δ and *uba4*Δ cells. **a**, Bulk tRNA. Note that the s^2U-modified tRNA migrates slightly faster through the gel than mcm^5s^2U-modified tRNA. **b**, Northern blot analysis using a specific probe against tE^{UUC}. **c**, Extracted ion chromatograms for protonated mcm^5s^2U nucleosides of single tRNA species tQ^{UUG} isolated from wild-type (red line) and *urm1*Δ (blue line) strains.

verified several of these interactions by yeast two-hybrid analysis, generating a *URM1* pathway protein–protein interaction network (Fig. 3a). These data support a role for Ncs2p, Ncs6p and Yor251cp in Urm1p function. Ncs2p and Ncs6p have previously been implicated in the protein modification function of Urm1p[6], whereas the function of *YOR251C* is unknown.

To verify these candidates functionally, we tested suppression of drug sensitivity by tRNA overexpression. Indeed, elevated tRNA levels rescued the growth phenotypes of *ncs2*Δ and *ncs6*Δ cells (Fig. 3b and Supplementary Fig. 2c). Surprisingly, this was not the case for *yor251c*Δ cells (data not shown), implying that Yor251cp may perform additional functions. APM gel and northern blot analyses showed that thiolated tRNAs were completely absent in the cytoplasm of *ncs2*Δ and *ncs6*Δ cells (Fig. 3c and Supplementary Fig. 2d), whereas they were significantly reduced in *yor251c*Δ cells (Fig. 3c and Supplementary Fig. 2d, e). Moreover, *YOR251C* forms a functional cluster with *URM1*, *NCS2*, *NCS6*, *UBA4* and the ELP complex genes in a large-scale SGA data set (M.C. and C.B., unpublished observations). Although *YOR251C* needs to be characterized, its domain organization is typical for mercaptopyruvate sulphurtransferases, suggesting that it may help in mobilizing the sulphur atom that will be attached to the tRNA.

To understand the molecular mechanism of tRNA thiolation better, we analysed the sequence of Ncs2p and Ncs6p. Ncs6p belongs to the TtcA protein family, which is important in the thiolation of cytidine[17]. *NCS6* and its homologues (Supplementary Fig. 6a) are essential for thiolation of uridine residues in tRNA of yeast, worms and *Thermus thermophilus*[18–20]. Interestingly, the *T. thermophilus* gene TtuA is preceded by a gene encoding for a small sulphur carrier with sequence similarity to Urm1p (Supplementary Fig. 6b), which participates in uridine thiolation. Additionally, the *NCS6* homologue in *Thermoanaerobacter tengcongensis* is found in a gene cluster with homologues of *URM1* and *UBA4*, further substantiating the idea of a common function.

Ncs6p contains a PP-loop ATPase domain, which is found in RNA modifying enzymes that bind the tRNA moiety and activate it through adenylation, like MnmA or ThiI[21,22]. Indeed, recombinant MBP-Ncs6p, but not MBP-Uba4p, directly bound yeast tRNA *in vitro* (Fig. 3d). Ncs2p shares some motives with Ncs6p and contains a divergent PP-loop ATPase domain (Supplementary Fig. 6a). However, because recombinant Ncs2p cloned to different tags was insoluble, we could not examine its molecular role *in vitro*.

All identified pathway members have homologues in worms and mammalian species (Supplementary Table 2). Thus, we compared by APM gel analysis tRNAs from wild-type (N2) and *tm1297* worms, which carry a deletion in the putative *NCS6* homologue. As tRNA tK^{UUU}, tE^{UUC} and tQ^{UUG} shifted in wild-type but not in *tm1297* mutants (Fig. 3e and data not shown), we named the ORF *tut-1* (thiolation of uridine in tRNA). Further, we established that MBP-TUT-1 directly binds to tRNA *in vitro* (Fig. 3d). Taken together, these

©2009 Macmillan Publishers Limited. All rights reserved

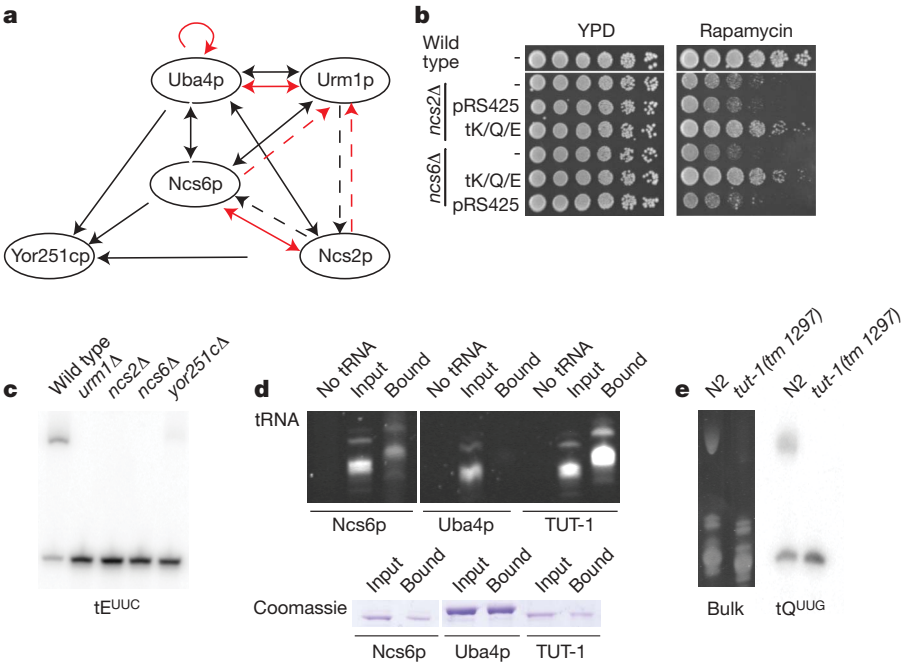


Figure 3 | Ncs2p, Ncs6p and Yor251cp function in the *URM1* pathway. **a**, *URM1* pathway protein–protein interaction detected by mass spectrometry (black lines) and yeast two-hybrid assay (red lines). Double head: interaction confirmed in the reciprocal experiment; dotted lines: interaction not confirmed in the reciprocal experiment. **b**, Spot dilution assays of *ncs2*Δ and *ncs6*Δ cells expressing elevated levels of tE^{UUC}, tK^{UUU} and tQ^{UUG} on YPD plates containing 2 nM rapamycin. **c**, Northern blot analysis of bulk tRNA from wild-type, *urm1*Δ, *ncs2*Δ, *ncs6*Δ and *yor251c*Δ cells using a specific probe for tE^{UUC}. **d**, *In vitro* binding assay with yeast tRNA and immobilized MBP-Ncs6p, MBP-Uba4p and MBP-TUT-1. tRNA gel (top). No bacterial tRNA was carried over from the purification. Proteins in the input and bound fractions were visualized by Coomassie blue staining (bottom). **e**, Analysis of bulk tRNA from wild-type (N2) and *tut-1*(*tm1297*) worms. APM gel (left) and northern blot analysis using a specific probe for tQ^{UUG} (right).

results demonstrate that tRNA thiolation is a conserved function of the *URM1 pathway* in *S. cerevisiae*, *Caenorhabditis elegans* and possibly all higher eukaryotes.

Our molecular analysis allows for two interpretations: either Urm1p acts as a sulphur carrier like MoaD and ThiS, or Urm1p is conjugated to another protein resulting in U_{34} thiolation. We did not identify Ahp1p, the only known Urm1p protein target[5], in our genetic screens, and tRNA from *ahp1*Δ cells showed wild-type tRNA thiolation levels (Supplementary Fig. 7a), excluding a role for Ahp1p urmylation in tRNA thiolation.

A strong argument in favour of Urm1p acting as a sulphur carrier stems from the domain structure of Uba4p (Supplementary Fig. 7b), which consists of an amino (N)-terminal nucleotide-binding domain (NBD), followed by a MoeZ motif and a carboxy (C)-terminal rhodanese domain (RHD). NBDs are required for activating MoaD and ThiS in bacteria, whereas RHDs containing proteins are responsible for transferring sulphur to the activated MoaD/ThiS proteins to form C-terminal thiocarboxylates[23–25]. Therefore Uba4p may activate Urm1p and convert it to a thiocarboxylate[26].

To study the role of Uba4p, we tagged it and mutated C225, the residue corresponding to the active site cysteine of UBLs E1-enzymes, and C397, the active cysteine of the RHD, to serine residues. Although the mutant proteins were expressed at wild-type levels (Supplementary Fig. 7c), they failed to complement the growth (Fig. 4a and Supplementary Fig. 7d) and tRNA thiolation (Fig. 4b and Supplementary Fig. 7e) defects of *uba4*Δ cells, demonstrating that both cysteine residues are critical for Uba4p function *in vivo*. Taken together, our findings reveal a crucial role for the Uba4p RHD, possibly by forming a thiocarboxylate intermediate on Urm1p.

To demonstrate unambiguously that Urm1p can act as a sulphur carrier, purified recombinant His_6-Urm1p and MBP-Uba4p were incubated in the presence of ATP and [^{35}S]cysteine. Sulphur was rapidly transferred from Uba4p to Urm1p (Fig. 4c and Supplementary Fig. 7f–h), and mass spectrometry combined with SUMmOn[27] analysis confirmed the presence of thiocarboxylate at the Urm1p C terminus (Supplementary Fig. 8). As these reactions did not contain Nfs1p, which is required *in vivo* as a sulphur donor[28], it is likely that our purified samples were contaminated with trace amounts of a cysteine desulphurase or with a sulphur derivative, like an oxidized form of ^{35}S. Interestingly, we observed an ATP-dependent appearance of a faster-migrating Urm1p form (Fig. 4c and Supplementary Fig. 7g), which most likely represents adenylated Urm1p, in analogy to MoaD[23]. We propose the following tRNA modification mechanism by the *URM1* pathway in eukaryotes (Fig. 4d): Urm1p, adenylated by the NBD of Uba4p, is converted to a Urm1p-thiocarboxylate by Uba4p RHD. The sulphur is then passed onto U_{34} either directly or with the help of Ncs6p/Ncs2p. Ncs6p binds tRNA and most likely adenylates U_{34}, thereby facilitating substitution of the carbonyl by a thiocarbonyl group. By forming a complex with Ncs6p, Ncs2p may influence the tRNA binding activity of Ncs6p or it might play an enzymatic role in the thiolation process. Yor251cp might contribute to a sulphur relay upstream of Uba4p; however, its precise role remains to be investigated.

tRNA modification appears to be one major function of Urm1p and independent of the protein conjugation pathway reported previously[4–6]. Our analysis suggests that the decrease in translation fidelity that accompanies loss of U_{34} thiolation might explain many of the *urm1*Δ phenotypes. Moreover, the presence of a thiocarboxylate at the C terminus of Urm1p raises the possibility that it might be conjugated to Ahp1p by a cysteine residue, like Pex5p, which is attached to ubiquitin by a dithiothreitol (DTT)-sensitive bond[29]. Thus, it will be interesting to analyse the chemical linkage of Urm1p to its protein targets, and to investigate how the thiolation and protein modification functions of Urm1p are related.

Our tRNA overexpression experiments functionally link Urm1p-mediated tRNA thiolation to oxidative stress, Tor-signalling and the cellular response to starvation. They suggest that tRNA levels, localization and modification state may play a regulatory role in this context. Finally, the identification of the eukaryotic system generating

©2009 Macmillan Publishers Limited. All rights reserved

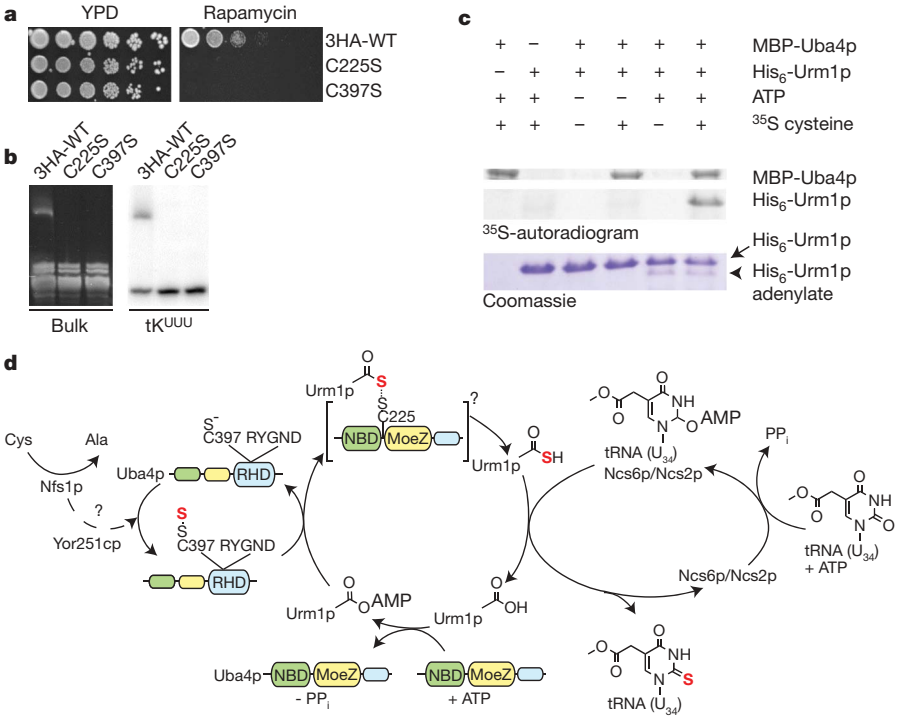


Figure 4 | Urm1p acts as a sulphur carrier. a, Spot dilution assay of 3HA-UBA4 cells and the indicated mutants on YPD plates containing 2 mM rapamycin. **b**, APM gel (left) and northern blot analyses of bulk tRNA isolated from 3HA-UBA4 and mutant yeast using a specific probe for tK^{UUU} (right). **c**, *In vitro* formation of adenylated and thiocarboxylated Urm1p. Purified MBP-Uba4p and His_6-Urm1p were included in the reaction as indicated and incubated with or without ATP and [^{35}S]cysteine. Upper panel: ^{35}S-autoradiogram; bottom panel: Coomassie staining. The adenylated form of Urm1p is marked by an arrowhead. **d**, A speculative model for Urm1p-mediated U_{34} thiolation. See main text for description.

thiolated uridine at the wobble position of cytoplasmic tRNA might also have medical implications, as the HIV reverse transcriptase relies on this modification to initiate reverse transcription of the HIV RNA genome *in vitro*[30].

METHODS SUMMARY

SGA was performed according to established protocols. tRNA was purified using column purification schemes designed for purification of bulk tRNA or single tRNA species. Proteins for *in vitro* assays were expressed and purified from *Escherichia coli*. Mass spectrometry procedures for RNA and protein samples are described in Supplementary Methods.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 10 March; accepted 14 November 2008.
Published online 14 January 2009.

1. Haglund, K. & Dikic, I. Ubiquitylation and cell signaling. *EMBO J.* **24,** 3353–3359 (2005).
2. Jentsch, S. & Pyrowolakis, G. Ubiquitin and its kin: how close are the family ties? *Trends Cell Biol.* **10,** 335–342 (2000).
3. Xu, J. *et al.* Solution structure of Urm1 and its implications for the origin of protein modifiers. *Proc. Natl Acad. Sci. USA* **103,** 11625–11630 (2006).
4. Furukawa, K., Mizushima, N., Noda, T. & Ohsumi, Y. A protein conjugation system in yeast with homology to biosynthetic enzyme reaction of prokaryotes. *J. Biol. Chem.* **275,** 7462–7465 (2000).
5. Goehring, A. S., Rivers, D. M. & Sprague, G. F. Jr. Attachment of the ubiquitin-related protein Urm1p to the antioxidant protein Ahp1p. *Eukaryot. Cell* **2,** 930–936 (2003).
6. Goehring, A. S., Rivers, D. M. & Sprague, G. F. Jr. Urmylation: a ubiquitin-like pathway that functions during invasive growth and budding in yeast. *Mol. Biol. Cell* **14,** 4329–4341 (2003).
7. Rubio-Texeira, M. Urmylation controls Nil1p and Gln3p-dependent expression of nitrogen-catabolite repressed genes in *Saccharomyces cerevisiae*. *FEBS Lett.* **581,** 541–550 (2007).
8. Jeong, J. S., Kwon, S. J., Kang, S. W., Rhee, S. G. & Kim, K. Purification and characterization of a second type thioredoxin peroxidase (type II TPx) from *Saccharomyces cerevisiae*. *Biochemistry* **38,** 776–783 (1999).
9. Huang, B., Lu, J. & Bystrom, A. S. A genome-wide screen identifies genes required for formation of the wobble nucleoside 5-methoxycarbonylmethyl-2-thiouridine in *Saccharomyces cerevisiae*. *RNA* **14,** 2183–2194 (2008).
10. Frohloff, F., Fichtner, L., Jablonowski, D., Breunig, K. D. & Schaffrath, R. *Saccharomyces cerevisiae* Elongator mutations confer resistance to the *Kluyveromyces lactis* zymocin. *EMBO J.* **20,** 1993–2003 (2001).
11. Fichtner, L. *et al.* Elongator's toxin-target (TOT) function is nuclear localization sequence dependent and suppressed by post-translational modification. *Mol. Microbiol.* **49,** 1297–1307 (2003).
12. Lu, J., Huang, B., Esberg, A., Johansson, M. J. & Bystrom, A. S. The *Kluyveromyces lactis* gamma-toxin targets tRNA anticodons. *RNA* **11,** 1648–1654 (2005).
13. Ashraf, S. S. *et al.* Single atom modification (O→S) of tRNA confers ribosome binding. *RNA* **5,** 188–194 (1999).
14. Esberg, A., Huang, B., Johansson, M. J. & Bystrom, A. S. Elevated levels of two tRNA species bypass the requirement for elongator complex in transcription and exocytosis. *Mol. Cell* **24,** 139–148 (2006).
15. Igloi, G. L. Interaction of tRNAs and of phosphorothioate-substituted nucleic acids with an organomercurial. Probing the chemical environment of thiolated residues by affinity electrophoresis. *Biochemistry* **27,** 3842–3849 (1988).
16. Umeda, N. *et al.* Mitochondria-specific RNA-modifying enzymes responsible for the biosynthesis of the wobble base in mitochondrial tRNAs. Implications for the molecular pathogenesis of human mitochondrial diseases. *J. Biol. Chem.* **280,** 1613–1624 (2005).
17. Jager, G., Leipuviene, R., Pollard, M. G., Qian, Q. & Bjork, G. R. The conserved Cys-X1–X2-Cys motif present in the TtcA protein is required for the thiolation of cytidine in position 32 of tRNA from *Salmonella enterica* serovar *Typhimurium*. *J. Bacteriol.* **186,** 750–757 (2004).
18. Bjork, G. R., Huang, B., Persson, O. P. & Bystrom, A. S. A conserved modified wobble nucleoside (mcm5s2U) in lysyl-tRNA is required for viability in yeast. *RNA* **13,** 1245–1255 (2007).
19. Dewez, M. *et al.* The conserved Wobble uridine tRNA thiolase Ctu1-Ctu2 is required to maintain genome integrity. *Proc. Natl Acad. Sci. USA* **105,** 5459–5464 (2008).
20. Shigi, N., Sakaguchi, Y., Suzuki, T. & Watanabe, K. Identification of two tRNA thiolation genes required for cell growth at extremely high temperatures. *J. Biol. Chem.* **281,** 14296–14306 (2006).
21. Kambampati, R. & Lauhon, C. T. MnmA and IscS are required for in vitro 2-thiouridine biosynthesis in *Escherichia coli*. *Biochemistry* **42,** 1109–1117 (2003).
22. Mueller, E. G. & Palenchar, P. M. Using genomic information to investigate the function of ThiI, an enzyme shared between thiamin and 4-thiouridine biosynthesis. *Protein Sci.* **8,** 2424–2427 (1999).
23. Leimkuhler, S., Wuebbens, M. M. & Rajagopalan, K. V. Characterization of *Escherichia coli* MoeB and its involvement in the activation of molybdopterin synthase for the biosynthesis of the molybdenum cofactor. *J. Biol. Chem.* **276,** 34695–34701 (2001).

©2009 Macmillan Publishers Limited. All rights reserved

24. Palenchar, P. M., Buck, C. J., Cheng, H., Larson, T. J. & Mueller, E. G. Evidence that ThiI, an enzyme shared between thiamin and 4-thiouridine biosynthesis, may be a sulfurtransferase that proceeds through a persulfide intermediate. *J. Biol. Chem.* **275,** 8283–8286 (2000).
25. Taylor, S. V. *et al.* Thiamin biosynthesis in *Escherichia coli*. Identification of this thiocarboxylate as the immediate sulfur donor in the thiazole formation. *J. Biol. Chem.* **273,** 16555–16560 (1998).
26. Schmitz, J. *et al.* The sulfurtransferase activity of Uba4 presents a link between ubiquitin-like protein conjugation and activation of sulfur carrier proteins. *Biochemistry* **47,** 6479–6489 (2008).
27. Pedrioli, P. G. *et al.* Automated identification of SUMOylation sites using mass spectrometry and SUMmOn pattern recognition software. *Nature Methods* **3,** 533–539 (2006).
28. Marelja, Z., Stocklein, W., Nimtz, M. & Leimkuhler, S. A novel role for human Nfs1 in the cytoplasm: Nfs1 acts as a sulfur donor for MOCS3, a protein involved in molybdenum cofactor biosynthesis. *J. Biol. Chem.* **283,** 25178–25185 (2008).
29. Williams, C., van den Berg, M., Sprenger, R. R. & Distel, B. A conserved cysteine is essential for Pex4p-dependent ubiquitination of the peroxisomal import receptor Pex5p. *J. Biol. Chem.* **282,** 22534–22543 (2007).
30. Isel, C. *et al.* Specific initiation and switch to elongation of human immunodeficiency virus type 1 reverse transcription require the post-transcriptional modifications of primer tRNA3Lys. *EMBO J.* **15,** 917–924 (1996).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank C. Rupp for technical support, M. Sohrmann for yeast strains, G. Rabut for the gift of the pGR51 plasmid, G. L. Igloi and K.-H. Altmann for providing APM, S. Mitani for *tut-1*(*tm1297*), B. Casado-Fernandez and D. Hilvert for discussion and technical advice, and T. Kurz and J. D. Alfonzo for reading the manuscript. This work was supported by grants from EuroDyna (to S.L.), the Roche Research Foundation (to P.G.A.P.), ETHZ (to S.L. and M.P.), the SNF (to M.P.), the Canadian Institutes of Health Research (to C.B.) and Genome Canada through the Ontario Genomics Institute (to C.B.).

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to M.P. (matthias.peter@bc.biol.ethz.ch) or K.H. (kay.hofmann@miltenyibiotec.de).

©2009 **Macmillan Publishers Limited. All rights reserved**

METHODS

Yeast and worm strains, plasmids and probes are listed in Supplementary Tables 3–5.

SGA analysis. SGA analysis was performed essentially as described[31–33]. Yeast growth was scored computationally, and overlapping hits in sister screens were considered as true positives. Statistical enrichment was calculated using the GOlorize plug-in for Cytoscape[34].

Immunoblotting and antibodies. The Immun-star (Biorad) detection kit was used, following the manufacturer's instructions. Antibodies were mouse anti-HA (C12A5), mouse anti-actin (Chemicon) and secondary HRP anti-mouse antibody (Biorad).

Generation of endogenously tagged yeast. Deletion strains (Euroscarf) were transformed with a plasmid system (by G. Rabut) allowing for the exchange of the G418 cassette by a tagged version of the gene of interest. Mutant *UBA4* was generated using the QuickChange Site-Directed Mutagenesis Kit (Stratagene).

Preparation of bulk tRNA and single tRNA species. Bulk tRNA was prepared as described[35], using Nucleobond PC100 columns (Machinery-Nagel). Single tRNA species were purified using the oligonucleotides listed in Supplementary Table 5 as described[36].

PAGE/APM-PAGE gel electrophoresis. Bulk (0.5 μg) or single (0.1 μg) tRNA per lane was analysed on 8% or 10% acrylamide gels; 0.5 × TBE; 7 M urea. APM was added to 50 μg ml^{-1} final concentration. Gels were stained with SYBR-gold (Invitrogen).

Northern blot analysis. Northern blot analysis was accomplished essentially as described[37], using the probes listed in Supplementary Table 5. Membranes were exposed to a phosphoimager plate (Molecular Dynamics) and scanned on a Storm 860 instrument (GE Healthcare).

Yeast drug sensitivity assays. Overnight cultures were diluted to $OD_{600\,nm} = 1$. Three microlitres of 1:5 serial dilutions were spotted on diamide[38], caffeine[39] and rapamycin[40] as indicated in the figure legends. Plates were imaged after three days at 30 °C. In some cases, yeast cells were grown at 37 °C to aggravate phenotypes.

Protein purification. Proteins were purified as described[41]. *E. coli* cultures were induced with 0.5 mM IPTG at 15 °C or 30 °C. Cell lysates were purified over His-NTA beads (Qiagen) or Amylose resin (New England Biolabs) according to the manufacturer's description. Protein samples were dialysed against 50 mM Hepes (pH 7.6), their purity and concentration controlled by SDS–PAGE and Bradford assay.

Urm1p *in vitro* thiolation and mass spectrometry. Eleven nanomoles of [^{35}S]cysteine (GE Healthcare or Perkin Elmer) was added to 22 pmol of MBP-Uba4p and 160 pmol His_6-Urm1p in 20 μl thiolation buffer (50 mM Hepes (pH 7.6), 50 mM KCl, 12 mM $Mg(OAc)_2$, 0.1 mM DTT, 2.5% glycerol, 40 μM PLP, 2 mM ATP) for 10 min at 30 °C and the reaction was stopped by adding non-reducing loading dye. The gel was fixed, dried and analysed on a phosphoimager. For the analysis of adenylated Urm1p, 0.1 units of Uba4p were added to 5 μg His_6-Urm1p and His_6-Urm1pGGΔ. For mass spectrometry, a reaction containing cold cysteine was used and analysed on an LTQ instrument. Mass spectrometry data was searched using SUMmOn[27] against an empty target protein database. The condition file specified the C-terminal tryptic peptide from Urm1p as a modification.

Time course experiments. Uba4p (0.1 units) and 5 μg of His_6-Urm1 or His_6-Urm1GGΔ were used. Thiolation buffer with [^{35}S]cysteine, but without DTT and PLP, was incubated at 30 °C. Samples were taken at the indicated time points and the reaction stopped by the addition of loading buffer. The signal was quantified using ImageJ.

tRNA binding assay. 500 pmol of proteins were incubated 1h on ice with 5 μg wild-type tRNA in binding buffer (50 mM Hepes (pH 7.6), 150 mM KOAc, 2 mM MgOAc, 0.25 mg ml^{-1} BSA, 0.001% Triton-X-100). MBP-beads (25 μl) were added and incubated for 1.5 h at 4 °C. The reactions were washed, split and mixed with either RNA or protein loading dye, respectively.

Nematodes. Handling of *C. elegans* was according to standard procedures[42]. To obtain tRNA, worms were lysed in Trizol (Invitrogen). Column purification was performed as for yeast tRNA.

RNA mass spectrometry. tRNA was digested and dephosphorylated into nucleosides as described[43]. The sample was bound to a hypercarb TopTip solid phase extraction column (Glygen), washed with 0.1% TFA in water, eluted with 0.1% TFA in 80% acetonitrile and lyophilized. Samples were resuspended in buffer A (0.1% formic acid in water), and separated on a 5-μm Hypercarb column (ThermoFinnigan) connected to an Agilent 1100 (Agilent Technologies). The HPLC used a linear gradient from 2 to 98% buffer B (100% acetonitrile) in 90 min at 60 °C. The column was connected to a LTQ-FT instrument (ThermoFinnigan). Raw data were converted to mzXML[44] format and analysed with the in-house-developed software AYB (P.G.A.P., unpublished). For further explanations, see Supplementary Fig. 3.

Protein immunoprecipitation and preparation for mass spectrometry. Yeast cultures were grown at 30 °C to $OD_{600\,nm}$ 1–1.5, pelleted, washed and lysed with glass beads in lysis buffer (50 mM Tris (pH 7.5), 250 mM NaCl, 5 mM $MgCl_2$, 0.5% NP-40, 50 mM NaF, 20 mM β-glycerophosphate) supplemented with complete EDTA-free Protease Inhibitor (Roche). The lysate was cleared and protein concentration adjusted to 10 mg ml^{-1}. The lysates were incubated for 3 h on a rotating wheel with either IgG-Sepharose (GE Healthcare) for TAP-tagged proteins, or monoclonal anti-HA-Agarose (Sigma) for HA-tagged proteins. Proteins were washed with lysis and wash buffer (20 mM Tris (pH 7.5), 100 mM NaCl, 5 mM $MgCl_2$, 5 mM NaF, 5 mM β-glycerophosphate). Proteins were eluted by 0.2 M Glycine (pH 2.5), neutralized with NH_4HCO_3 (pH 8), reduced for 30 min with 5 mM DTT at 60 °C and carboxyamidomethylated using 15 mM iodoacetamide for 45 min at room temperature in the dark. The proteins were digested overnight at 37 °C with sequencing-grade modified Trypsin (Promega). Peptides were purified on C18 MicroSpin Columns (The Nest Group) before mass spectrometry analysis.

Protein mass spectrometry analysis. Peptides were separated on a 10 cm × 75 μm C18 fused-silica column online with an LTQ instrument (ThermoFinnigan). After conversion to the mzXML format, mass spectrometry data were searched against the National Cancer Institute yeast protein database using the X!Tandem[45] with the K-score plug-in[46]. Search parameters allowed for a variable modification of 15.99 daltons (Da) at methionine and of 57.06 Da at cysteine residues. Peptide and protein probabilities were evaluated with PeptideProphet[47] and ProteinProphet[48], respectively.

31. Tong, A. H. *et al.* Systematic genetic analysis with ordered arrays of yeast deletion mutants. *Science* **294,** 2364–2368 (2001).
32. Tong, A. H. *et al.* Global mapping of the yeast genetic interaction network. *Science* **303,** 808–813 (2004).
33. Tong, A. H. Y., Boone, C., Stansfield, I. & Michael, J. R. S. 16 High-throughput strain construction and systematic synthetic lethal screening. *Meths Microbiol.* **36,** 369–386, 706–707 (2007).
34. Garcia, O. *et al.* GOlorize: a Cytoscape plug-in for network visualization with Gene Ontology-based layout and coloring. *Bioinformatics* **23,** 394–396 (2007).
35. Bjork, G. R. *et al.* A primordial tRNA modification required for the evolution of life? *EMBO J.* **20,** 231–239 (2001).
36. Suzuki, T. & Suzuki, T. Chaplet column chromatography: isolation of a large set of individual RNAs in a single step. *Meths Enzymol.* **425,** 231–239 (2007).
37. Lau, N. C., Lim, L. P., Weinstein, E. G. & Bartel, D. P. An abundant class of tiny RNAs with probable regulatory roles in *Caenorhabditis elegans*. *Science* **294,** 858–862 (2001).
38. Kosower, N. S., Kosower, E. M., Wertheim, B. & Correa, W. S. Diamide, a new reagent for the intracellular oxidation of glutathione to the disulfide. *Biochem. Biophys. Res. Commun.* **37,** 593–596 (1969).
39. Wanke, V. *et al.* Caffeine extends yeast lifespan by targeting TORC1. *Mol. Microbiol.* **69,** 277–285 (2008).
40. Kunz, J. *et al.* Target of rapamycin in yeast, TOR2, is an essential phosphatidylinositol kinase homolog required for G1 progression. *Cell* **73,** 585–596 (1993).
41. Braun, P. *et al.* Proteome-scale purification of human proteins from bacteria. *Proc. Natl Acad. Sci. USA* **99,** 2654–2659 (2002).
42. Brenner, S. The genetics of *Caenorhabditis elegans*. *Genetics* **77,** 71–94 (1974).
43. Gehrke, C. W., Kuo, K. C., McCune, R. A., Gerhardt, K. O. & Agris, P. F. Quantitative enzymatic hydrolysis of tRNAs: reversed-phase high-performance liquid chromatography of tRNA nucleosides. *J. Chromatogr.* **230,** 297–308 (1982).
44. Pedrioli, P. G. *et al.* A common open representation of mass spectrometry data and its application to proteomics research. *Nature Biotechnol.* **22,** 1459–1466 (2004).
45. Craig, R. & Beavis, R. C. TANDEM: matching proteins with tandem mass spectra. *Bioinformatics* **20,** 1466–1467 (2004).
46. MacLean, B., Eng, J. K., Beavis, R. C. & McIntosh, M. General framework for developing and evaluating database scoring algorithms using the TANDEM search engine. *Bioinformatics* **22,** 2830–2832 (2006).
47. Nesvizhskii, A. I., Keller, A., Kolker, E. & Aebersold, R. A statistical model for identifying proteins by tandem mass spectrometry. *Anal. Chem.* **75,** 4646–4658 (2003).
48. Keller, A., Nesvizhskii, A. I., Kolker, E. & Aebersold, R. Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. *Anal. Chem.* **74,** 5383–5392 (2002).

©2009 Macmillan Publishers Limited. All rights reserved

LETTERS

Coenzyme recognition and gene regulation by a flavin mononucleotide riboswitch

Alexander Serganov[1]*, Lili Huang[1]* & Dinshaw J. Patel[1]

The biosynthesis of several protein cofactors is subject to feedback regulation by riboswitches[1–3]. Flavin mononucleotide (FMN)-specific riboswitches[4,5], also known as RFN elements[6], direct expression of bacterial genes involved in the biosynthesis and transport of riboflavin (vitamin B_2) and related compounds. Here we present the crystal structures of the *Fusobacterium nucleatum* riboswitch bound to FMN, riboflavin and antibiotic roseoflavin[7]. The FMN riboswitch structure, centred on an FMN-bound six-stem junction, does not fold by collinear stacking of adjacent helices, typical for folding of large RNAs. Rather, it adopts a butterfly-like scaffold, stapled together by opposingly directed but nearly identically folded peripheral domains. FMN is positioned asymmetrically within the junctional site and is specifically bound to RNA through interactions with the isoalloxazine ring chromophore and direct and Mg^{2+}-mediated contacts with the phosphate moiety. Our structural data, complemented by binding and footprinting experiments, imply a largely pre-folded tertiary RNA architecture and FMN recognition mediated by conformational transitions within the junctional binding pocket. The inherent plasticity of the FMN-binding pocket and the availability of large openings make the riboswitch an attractive target for structure-based design of FMN-like antimicrobial compounds. Our studies also explain the effects of spontaneous and antibiotic-induced deregulatory mutations and provided molecular insights into FMN-based control of gene expression in normal and riboflavin-overproducing bacterial strains.

Riboswitches that target protein cofactors constitute the most abundant class of untranslated messenger RNA (mRNA) regions capable of direct sensing of cellular metabolites[8]. Recent three-dimensional structures of thiamine pyrophosphate (TPP)[9–11] and *S*-adenosyl-L-methionine (SAM)[12,13] riboswitches provided the first insights into the molecular organization of cofactor-specific riboswitches. Nevertheless, additional structural studies are required to explain the principles of RNA folding and ligand recognition for larger riboswitches, such as those that target adocobalamin[14] and FMN[4,5]. The determination of the FMN riboswitch structure could be especially interesting and timely, given the predicted complexity of its riboswitch fold, the abundance of this riboswitch in pathogenic species[7,8] and its involvement in riboflavin overproduction in biotechnologically relevant bacterial strains[15].

Fusobacterium nucleatum plays a role in periodontal disease and other human infections and is considered one of the most pathogenic bacteria of the genus[16]. The intracellular concentration of FMN in *F. nucleatum* is apparently controlled by a transcription attenuation mechanism involving two riboswitches, positioned before the riboflavin synthetic genes of the *ribHDE(B/A)* operon and the candidate riboflavin transporter *impX* gene (Supplementary Fig. 1)[17].

We have determined the 2.95-Å structure of the FMN-bound 112-nucleotide *F. nucleatum impX* RFN element, which conforms well with the consensus sequence and the six-helical junctional secondary structure characteristic of this riboswitch family (Fig. 1a–c and Supplementary Fig. 2a, b). The structure features a complex FMN-bound junctional region stapled together by two peripheral domains, P2–P6 and P3–P5. Each peripheral domain is formed by two interacting stem-loops, stabilized by two pairs of tertiary contacts involving loop–loop (L2–L6 and L3–L5) and loop–helix (L6–P2 and L3–P5) interactions (Fig. 1a and Supplementary Fig. 3a, b), resulting in an overall butterfly-like fold (Fig. 1c). Stems P1 and P4 radiate in opposite directions from the bottom part of the junction. Most unexpectedly, the peripheral domains show nearly identical conformations when superposed by ~180° rotation along the axis directed through the central region of the riboswitch (Fig. 1d). Though symmetry is not a characteristic feature of large RNAs, such local symmetry has also been identified in the ribosome[18]. Surprisingly, the pseudo-symmetrical FMN does not take advantage of the twofold symmetry between riboswitch elements, because its isoalloxazine ring and phosphate are oriented towards different riboswitch domains. The peripheral domains contain small RNA motifs[19,20], such as a T-loop[8,21,22] (Supplementary Fig. 3c), and show a striking resemblance to larger architectural modules found in 23S ribosomal RNA (rRNA)[19] (Supplementary Fig. 4). Despite this similarity, the domains serve as molecular staples to hold and shape the FMN-binding pocket in the riboswitch, whereas in rRNA they form stable platforms for interactions with proteins and other rRNA segments[23].

Comparison of the loop–helix interactions uncovers a small difference that may contribute to gene expression regulation. In the P2–P6 domain, invariant G12 from the G12•(G93–C30) triple (Fig. 1e) replaces the corresponding A63 in the stable A-minor A63•(G41–C82) triple from the P3–P5 domain (Fig. 1f). To prevent a steric clash with the G93–C30 pair, the G12 base is moved out of the triple plane, thereby weakening both G12•(G93–C30) and G11•(G84–C31) triples. The lower stability of these adjacent triples could enhance the mobility of the J1-2 segment, highlighting its contribution to the coenzyme-sensitive switch governing gene regulation. In essence, the J1-2 segment participates in anti-terminator formation in the absence of FMN, whereas in the FMN-bound state it is locked up in the junction, thereby facilitating formation of the regulatory P1 helix.

In contrast to other multi-stem junctional riboswitches[9–11,13,24–27], the junctional region of the FMN riboswitch is not constructed on the basis of collinear stacking of adjacent helices, but instead is composed of several non-paired segments, which provide a smooth transition between adjacent helices (Fig. 1c and Supplementary Fig. 3d, e). Unlike coenzymes in other riboswitches[9–11,13], the oxidized FMN is positioned centrally inside the junctional region and is surrounded by all six RNA stems (Figs 1c and 2a). The planar isoalloxazine ring system intercalates between A48 and A85, thereby providing a continuous stacking

[1]Structural Biology Program, Memorial Sloan-Kettering Cancer Center, New York, New York 10065, USA.
*These authors contributed equally to this work.

©2009 Macmillan Publishers Limited. All rights reserved

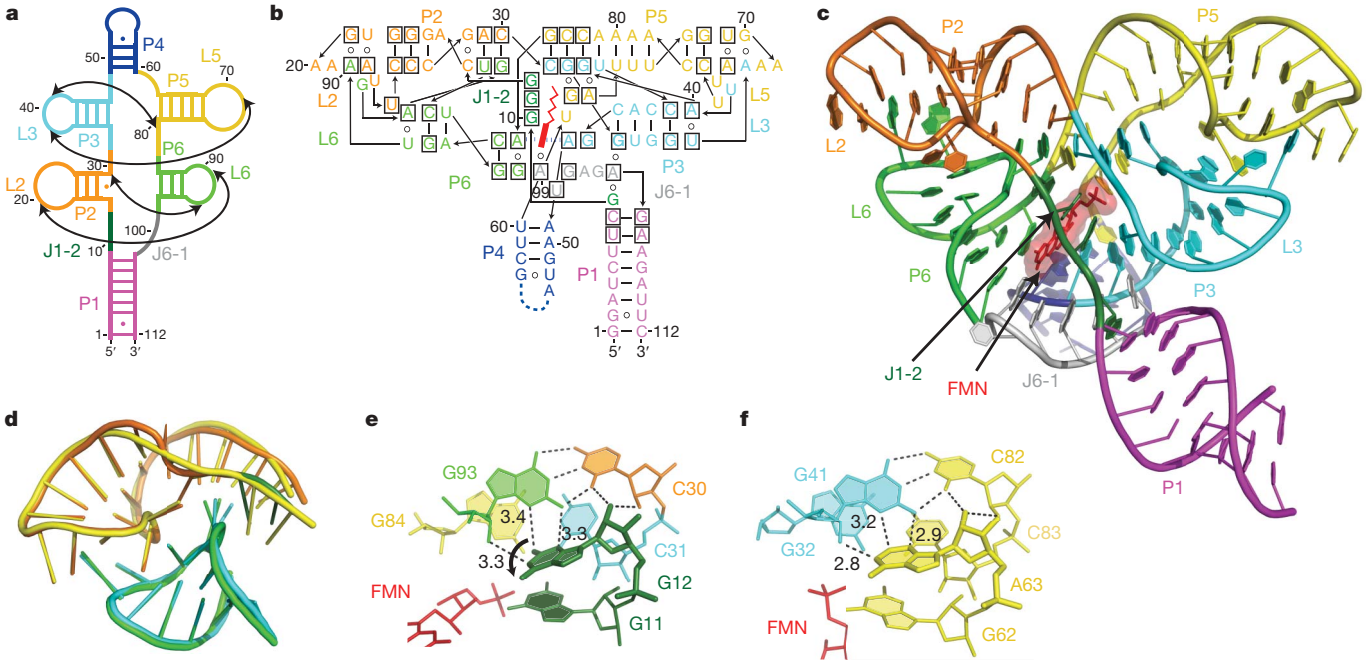


Figure 1 | Overall structure and tertiary interactions of the FMN-bound *F. nucleatum* riboswitch. **a**, Homology-based schematic of the FMN riboswitch with key long-range interactions indicated by arrows. RNA segments are depicted in colours used for subsequent figures. **b**, Schematic of the riboswitch fold observed in the crystal structure of the complex. The bound FMN is in red. Key stacking interactions involving FMN are shown as blue dashed lines. Nucleotides that are more than 95% conserved among 183 FMN riboswitches are boxed. **c**, Overall riboswitch structure in a ribbon representation. **d**, Superposition of the P2–P6 (nucleotides 10–32 and 85–98) and P3–P5 domains (nucleotides 62–84 and 33–46). The root mean square deviation is 1.8 Å. **e**, **f**, Distinct alignments of nucleotide triples in the P2–P6 (**e**) and P3–P5 (**f**) domains. Dashed lines depict putative hydrogen bonds. Distances are in ångstroms.

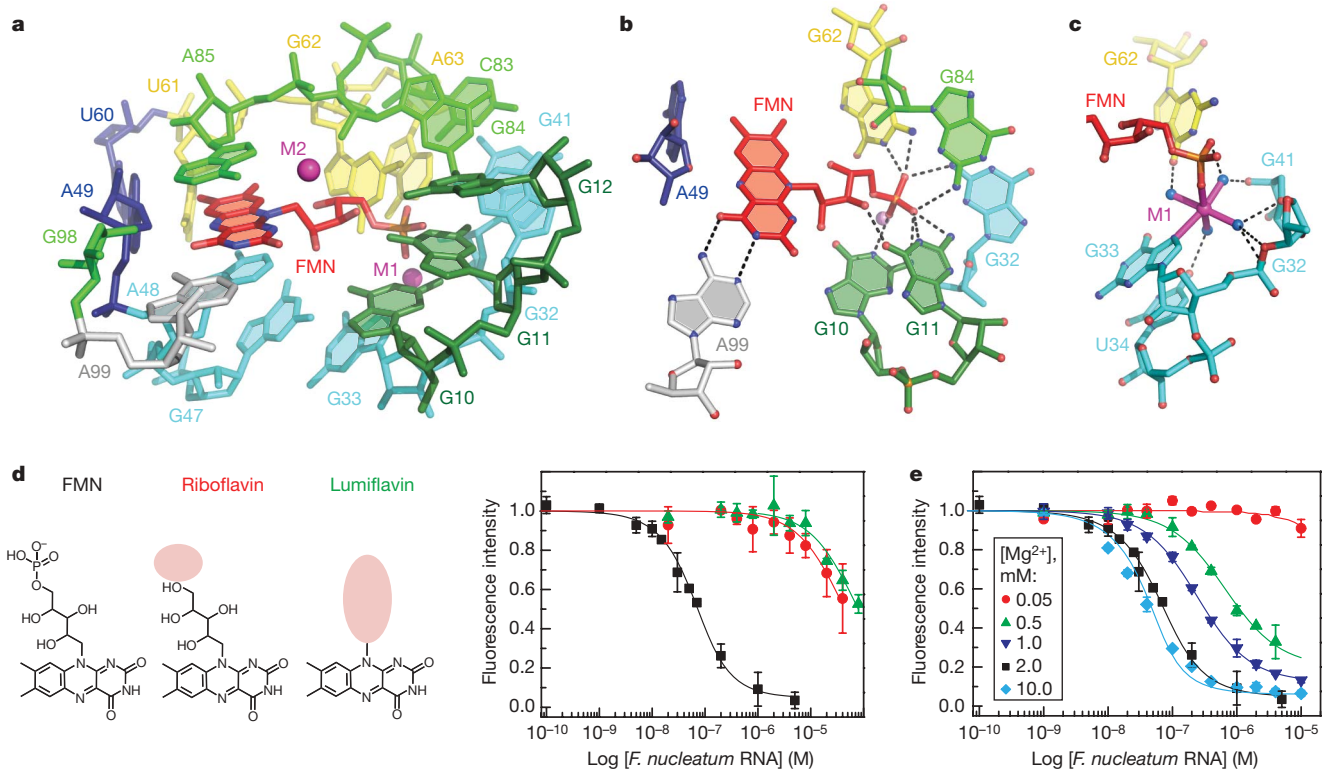


Figure 2 | Recognition of FMN by its riboswitch. **a**, A view of the junction with bound FMN. Magenta spheres depict cations in proximity to FMN. **b**, Details of riboswitch–FMN interactions. Hydrogen bond distances are shown with dashed lines. **c**, FMN–riboswitch interactions mediated by M1, assigned to an Mg^{2+} cation. Magenta sticks depict coordination bonds. Mg^{2+}-coordinated water molecules (blue spheres) are positioned on the basis of direct coordination bonds with phosphate and G33. **d**, Chemical structures (left panel) and fluorescent binding assay (right panel) of FMN and its analogues with the 112-nucleotide *F. nucleatum* riboswitch conducted in 2 mM $MgCl_2$ and 100 mM KCl. The dissociation constants (micromolar, mean ± s.d.) are: FMN, 0.0375 ± 0.0031; riboflavin, ~39.8; lumiflavin, ~59.0. **e**, Interaction of FMN with *F. nucleatum* riboswitch in 100 mM KCl as a function of variable Mg^{2+} concentration. The dissociation constants (nanomolar, mean ± s.d.) are: 0.5 mM $MgCl_2$, 406.0 ± 52.7; 1.0 mM $MgCl_2$, 211.7 ± 2.1; 10.0 mM $MgCl_2$, 10.9 ± 2.4.

©2009 Macmillan Publishers Limited. All rights reserved

alignment linking P6 and P3 helices. The uracil-like edge of the ring system forms specific Watson–Crick-like hydrogen bonds with the highly conserved A99 (Fig. 2b). The ribityl moiety of FMN uses only one of its four oxygens for hydrogen bonding, whereas phosphate oxygens form additional hydrogen bonds with Watson–Crick edges of several conserved guanines. The interaction between the phosphate of FMN and RNA is also bridged by metal ion M1, assigned as Mg^{2+}, which directly coordinates the phosphate oxygen of FMN and the N7 position of G33, and forms several water-mediated contacts with neighbouring nucleotides (Fig. 2c). In support of the structural data, the binding affinity of riboflavin, the FMN precursor, which lacks the phosphate and does not control gene expression, decreases by about 1,000-fold, compared with binding of FMN to both *F. nucleatum* and *Bacillus subtilis*[5] riboswitches (Fig. 2d and Supplementary Fig. 5). The additional removal of the ribityl moiety in lumiflavin reduces binding affinity to only a very small extent (Fig. 2d). In contrast to the FMN riboswitch, the *in-vitro*-selected RNA aptamer does not bind the phosphate moiety of FMN and uses the Hoogsteen edge of adenine for specific recognition of the ring system (Supplementary Fig. 6)[28].

The identity of metal M1 as Mg^{2+} has been confirmed by substitution with Mn^{2+}, a mimic of Mg^{2+} (Supplementary Fig. 7). Cs^{+} cations failed to replace M1 and were found in the P3 stem and next to the FMN ring system (Supplementary Fig. 8), where a Cs^{+} replaces a K^{+} (labelled M2, Fig. 2a) cation. The interactions between FMN and the riboswitch critically depend on the physiological concentration of Mg^{2+} (Fig. 2e and Supplementary Fig. 9a), and can be enhanced further by addition of 100 mM K^{+} (Supplementary Fig. 9b). Smaller and larger monovalent cations, such as Na^{+} and Cs^{+}, potentiate FMN binding to a lesser extent than K^{+} (Supplementary Fig. 9c). The identity of the divalent cation appears not to be crucial for FMN binding, because Mg^{2+} can be substituted by Ca^{2+}, as well as Ba^{2+} and Mn^{2+}, even in the absence of monovalent cations (Supplementary Figs 9a, b and 10). It is likely that cation M1 contributes significantly to the metal dependence of the FMN binding, because the coenzyme binds the riboswitch equally well in the presence of cobalt hexamine, which can be accommodated with some adjustments in the M1 position (Supplementary Figs 9b and 11), whereas bulkier iridium hexamine, which does not fit into the M1 position, reduces binding affinity about 15-fold (Supplementary Figs 9b and 12). Unlike the M1 site, the M2 position constitutes part of a larger cation-binding area, which extends towards G11 and FMN and which can accommodate all soaked metals, except Mn^{2+}. Though cations in this area stabilize the FMN-binding pocket, the weaker dependence of FMN binding on monovalent cations (Supplementary Fig. 9a, b) suggests a less critical role of the M2 site for FMN–riboswitch complex formation.

To gain further insights into the ligand discrimination by FMN riboswitches, we have determined crystal structures of the riboswitch in complex with riboflavin and roseoflavin[7] (Fig. 3a, b). In the riboflavin-bound structure, the P4 helix is shifted towards the ligand-binding pocket (Supplementary Fig. 13a) and several nucleotides that are close to the phosphate-binding cavity and the hydrophobic edge of the ring system become slightly re-positioned (Fig. 3a). The roseoflavin-bound structure adopts a conformation similar to the riboflavin-bound structure (Supplementary Fig. 13b), with additional spatial adjustments of U61 and G62 to accommodate the dimethylamino group that substitutes for a methyl on the hydrophobic edge of the isoalloxazine ring system (Fig. 3b). Such ligand-dependent conformational changes have been described earlier only for the TPP riboswitch[11]. In contrast to similar recognition of the ring system, the ribityl moieties of all three FMN riboswitch ligands adopt slightly different conformations and exhibit somewhat different interactions with the RNA. As expected, no metal M1 was found in the two analogue complexes, given that analogues lack a phosphate group. The transcription of the mRNA that contains the FMN riboswitch was found to be inhibited *in vitro* by another coenzyme, flavin adenine dinucleotide (FAD)[4], used at a 17-fold higher concentration than FMN. Our structure-based modelling of a FAD-riboswitch complex supports the possibility of FAD-mediated control of the FMN riboswitch[4] because FAD can be accommodated within the binding pocket after minor conformational adjustments (Supplementary Fig. 14). The observed plasticity of the ligand-binding pocket, together with the large openings next to the edges of the ring system (Fig. 3c) and the phosphate-binding site (Supplementary Fig. 15), makes the FMN riboswitch an attractive target for the structure-based design of FMN-like antimicrobial compounds.

Because the bound FMN is enveloped by the RNA, the FMN riboswitch is anticipated to re-arrange its conformation on complex formation. To access these potential conformational changes, we have performed footprinting experiments using nucleases V1 (paired and stacked regions) and T2 (single-stranded regions) (Fig. 4a). Despite similar conformations of the peripheral domains, the P2–P6 domain is more accessible to both nucleases (Supplementary Fig. 16), and exhibits more pronounced FMN-induced changes (Fig. 4b) than its P3–P5 counterpart, implicative of a more rigid conformation for the P3–P5 domain. Nevertheless, the FMN-induced changes in the peripheral domains are not extensive and are consistent with a largely pre-formed conformation for these regions. The presence of V1 and the absence of T2 cleavages indicate that nucleotides of the junctional segments are most likely involved in stacking interactions in the unbound state. Most of these nucleotides, which are clustered around the phosphate (nucleotides 10, 29–33) and isoalloxazine ring system (nucleotides 46–50, 97–102) of FMN and are adjacent to the P1 helix (nucleotides 9, 104), become protected on FMN binding owing to the mobility restrictions and shielding by neighbouring

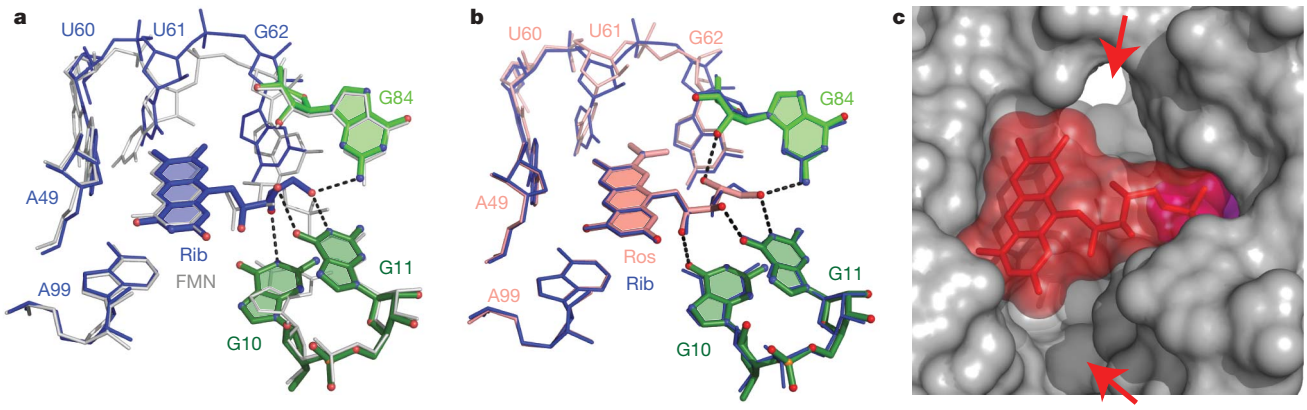


Figure 3 | Interactions of FMN analogues with the riboswitch. **a**, All-atom superposition of the ligand-binding pocket for riboflavin-bound (blue and green) and FMN-bound (grey) riboswitches. Nucleotides in green are positioned within hydrogen-bond distances of the ribityl moiety of riboflavin. **b**, Superposition of riboflavin-bound (blue) and roseoflavin-bound (pink and green) riboswitches, depicted as in **a**. **c**, Surface view inside of the FMN-bound riboswitch with large openings shown with red arrows.

©2009 Macmillan Publishers Limited. All rights reserved

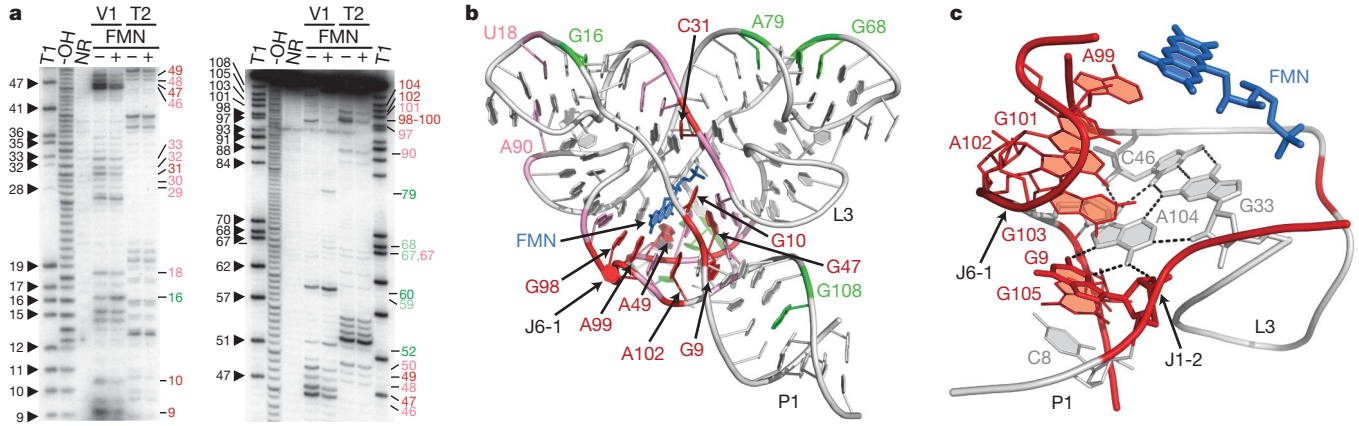


Figure 4 | Probing the FMN riboswitch tertiary structure. **a**, Probing of 5′ ^{32}P-labelled 112-nucleotide RNA (0.5 μM) by V1 and T2 nucleases in the absence and presence of sixfold excess of FMN. T1 and $^{-}$OH designate RNase T1 and alkaline ladders, respectively. NR, no reaction. Weak and strong FMN-induced cleavage protections are shown in light and dark red, whereas weak and strong cleavage enhancements are shown in light and dark green. **b**, Projections of the nuclease cleavage reductions (light and dark red) and enhancement (green) on the riboswitch structure. **c**, Nucleotides potentially facilitating formation of the regulatory P1 helix. Hydrogen bonds in the G9•A104•(G33–C46) tetrad are indicted. FMN-induced cleavage reductions in the in-line probing assay[5] are in red.

RNA segments. Our structural and footprinting data, corroborated by the in-line probing results on the *B. subtilis* riboswitch[5] (Supplementary Fig. 17), suggest that the primary recognition event(s) may involve Mg^{2+}-mediated interaction of the FMN phosphate with RNA, coupled with intercalation and base-specific interactions involving the isoalloxazine ring system. These interactions should stabilize the J1-2 segment (G10–G12) and, in accordance with the strongest protections (Fig. 4a, b), bring the nuclease-accessible J6-1 segment (A99–G104) and the C8–G9 step close to each other. Such positioning provides partial stacking between the J6-1 segment (A99–G104) and the helix P1-forming strands, and facilitates formation of the non-canonical G10•G47 pair, the type I A-minor motif-based G9•A104•(G33–C46) tetrad, the C8–G105 base pair, and other base pairs of the P1 helix (Fig. 4c).

Because riboflavin biosynthetic capability is lacking in higher animals, riboflavin is traditionally used for food and feed fortification. Riboflavin can be produced in bacterial strains selected as roseoflavin-resistant mutants with deregulated riboflavin biosynthesis[15]. The FMN riboswitch structure (Supplementary Fig. 18) readily explains the effects of deregulated mutations[15,29]. Surprisingly, most of the mutations are concentrated in the peripheral domains, where they destabilize the helices or directly disrupt tertiary contacts. Mutations G32A/C and G62A found in the FMN-binding pocket disrupt the G62•(C83–G32) triple (Fig. 1f), which interacts with the Mg^{2+}-coordinated phosphate of FMN. Two other mutations, G105U and G108A, prevent formation and/or affect stability of the regulatory P1 helix.

Recent studies have demonstrated slow kinetics of association and dissociation for the FMN–riboswitch complex, supportive of a kinetically driven riboswitch mechanism[30]. These kinetic characteristics are consistent with the recognition principles identified in our three-dimensional structure. Indeed, riboswitch folding requires formation of multiple non-canonical and tertiary interactions and other conformational adjustments on FMN binding, which together may account for the slow association rate. Subsequent FMN release is likely to be slowed down due to envelopment of the ligand by the RNA. The propensity of other large riboswitches to function as kinetically driven genetic switches remains a challenging area for future exploration.

METHODS SUMMARY

Crystallization and structure determination. The complexes of the riboswitch with FMN and its analogues were prepared by mixing 0.4 mM RNA with 0.7 mM ligand in a buffer containing 100 mM potassium acetate, pH 6.8, and 4 mM $MgCl_2$. FMN–riboswitch crystals were grown by hanging-drop vapour diffusion after mixing the complex with the reservoir solution (0.1 M MES-sodium, pH 6.5, 100 mM $MgCl_2$ and 10% (w/v) PEG 4000) at equimolar ratio. For soaking, crystals were incubated for 10 h in the reservoir solution supplemented with 25% glycerol and the following heavy atom salts: 5 mM $[Ir(NH_3)_6]Cl_3$, 15 mM $MnCl_2$, 10 mM $[Co(NH_3)_6]Cl_3$, 30 mM $BaCl_2$ and 30 mM CsCl. Crystals were flash frozen in liquid nitrogen and data were collected at 100 K. The structure was determined using 3.0-Å multiwavelength anomalous dispersion (MAD) iridium data (Supplementary Fig. 19 and Supplementary Table 1) and refined to R_{work}/R_{free} 20.0/24.3 with a native data set. FMN and cations were added to the model based on the analysis of $2F_o - F_c$, $F_o - F_c$ and anomalous electron density maps. Analogue-bound and metal-soaked riboswitch structures were refined using the native riboswitch model (Supplementary Table 2). Figures were prepared with PyMOL (http://www.pymol.org).

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 23 August; accepted 13 November 2008.
Published online 25 January 2009.

1. Nudler, E. & Mironov, A. S. The riboswitch control of bacterial metabolism. *Trends Biochem. Sci.* **29,** 11–17 (2004).
2. Serganov, A. & Patel, D. J. Ribozymes, riboswitches and beyond: regulation of gene expression without proteins. *Nature Rev. Genet.* **8,** 776–790 (2007).
3. Winkler, W. C. & Breaker, R. R. Regulation of bacterial gene expression by riboswitches. *Annu. Rev. Microbiol.* **59,** 487–517 (2005).
4. Mironov, A. S. *et al.* Sensing small molecules by nascent RNA: a mechanism to control transcription in bacteria. *Cell* **111,** 747–756 (2002).
5. Winkler, W. C., Cohen-Chalamish, S. & Breaker, R. R. An mRNA structure that controls gene expression by binding FMN. *Proc. Natl Acad. Sci. USA* **99,** 15908–15913 (2002).
6. Gelfand, M. S., Mironov, A. A., Jomantas, J., Kozlov, Y. I. & Perumov, D. A. A conserved RNA structure element involved in the regulation of bacterial riboflavin synthesis genes. *Trends Genet.* **15,** 439–442 (1999).
7. Blount, K. F. & Breaker, R. R. Riboswitches as antibacterial drug targets. *Nature Biotechnol.* **24,** 1558–1564 (2006).
8. Barrick, J. E. & Breaker, R. R. The distributions, mechanisms, and structures of metabolite-binding riboswitches. *Genome Biol.* **8,** R239 (2007).
9. Serganov, A., Polonskaia, A., Phan, A. T., Breaker, R. R. & Patel, D. J. Structural basis for gene regulation by a thiamine pyrophosphate-sensing riboswitch. *Nature* **441,** 1167–1171 (2006).
10. Thore, S., Leibundgut, M. & Ban, N. Structure of the eukaryotic thiamine pyrophosphate riboswitch with its regulatory ligand. *Science* **312,** 1208–1211 (2006).
11. Edwards, T. E. & Ferre-D'Amare, A. R. Crystal structures of the thi-box riboswitch bound to thiamine pyrophosphate analogs reveal adaptive RNA-small molecule recognition. *Structure* **14,** 1459–1468 (2006).
12. Gilbert, S. D., Rambo, R. P., Van Tyne, D. & Batey, R. T. Structure of the SAM-II riboswitch bound to *S*-adenosylmethionine. *Nature Struct. Mol. Biol.* **15,** 177–182 (2008).
13. Montange, R. K. & Batey, R. T. Structure of the *S*-adenosylmethionine riboswitch regulatory mRNA element. *Nature* **441,** 1172–1175 (2006).
14. Nahvi, A. *et al.* Genetic control by a metabolite binding mRNA. *Chem. Biol.* **9,** 1043–1049 (2002).

©2009 Macmillan Publishers Limited. All rights reserved

15. Burgess, C. M., Smid, E. J., Rutten, G. & van Sinderen, D. A general method for selection of riboflavin-overproducing food grade micro-organisms. *Microb. Cell Fact.* **5,** 24 (2006).
16. Kapatral, V. *et al.* Genome sequence and analysis of the oral bacterium *Fusobacterium nucleatum* strain ATCC 25586. *J. Bacteriol.* **184,** 2005–2018 (2002).
17. Vitreschak, A. G., Rodionov, D. A., Mironov, A. A. & Gelfand, M. S. Regulation of riboflavin biosynthesis and transport genes in bacteria by transcriptional and translational attenuation. *Nucleic Acids Res.* **30,** 3141–3151 (2002).
18. Agmon, I., Bashan, A., Zarivach, R. & Yonath, A. Symmetry at the active site of the ribosome: structural and functional implications. *Biol. Chem.* **386,** 833–844 (2005).
19. Jaeger, L., Verzemnieks, E. J. & Geary, C. The UA_handle: a versatile submotif in stable RNA architectures. *Nucleic Acids Res.* doi:10.1093/nar/gkn911 (28 November 2008).
20. Lescoute, A. & Westhof, E. The interaction networks of structured RNAs. *Nucleic Acids Res.* **34,** 6587–6604 (2006).
21. Krasilnikov, A. S. & Mondragon, A. On the occurrence of the T-loop RNA folding motif in large RNA molecules. *RNA* **9,** 640–643 (2008).
22. Nagaswamy, U. & Fox, G. E. Frequent occurrence of the T-loop RNA folding motif in ribosomal RNAs. *RNA* **8,** 1112–1119 (2002).
23. Selmer, M. *et al.* Structure of the 70S ribosome complexed with mRNA and tRNA. *Science* **313,** 1935–1942 (2006).
24. Batey, R. T., Gilbert, S. D. & Montange, R. K. Structure of a natural guanine-responsive riboswitch complexed with the metabolite hypoxanthine. *Nature* **432,** 411–415 (2004).
25. Serganov, A. *et al.* Structural basis for discriminative regulation of gene expression by adenine- and guanine-sensing mRNAs. *Chem. Biol.* **11,** 1729–1741 (2004).
26. Garst, A. D., Heroux, A., Rambo, R. P. & Batey, R. T. Crystal structure of the lysine riboswitch regulatory mRNA element. *J. Biol. Chem.* **283,** 22347–22351 (2008).
27. Serganov, A., Huang, L. & Patel, D. J. Structural insights into amino acid binding and gene control by a lysine riboswitch. *Nature* **455,** 1263–1267 (2008).
28. Fan, P., Suri, A. K., Fiala, R., Live, D. & Patel, D. J. Molecular recognition in the FMN-RNA aptamer complex. *J. Mol. Biol.* **258,** 480–500 (1996).
29. Burgess, C., O'Connell-Motherway, M., Sybesma, W., Hugenholtz, J. & van Sinderen, D. Riboflavin production in *Lactococcus lactis*: potential for *in situ* production of vitamin-enriched foods. *Appl. Environ. Microbiol.* **70,** 5769–5777 (2004).
30. Wickiser, J. K., Winkler, W. C., Breaker, R. & Crothers, D. M. The speed of RNA transcription and metabolite binding kinetics operate an FMN riboswitch. *Mol. Cell* **18,** 49–60 (2005).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank the personnel of beamline X29 at the Brookhaven National Laboratory and beamline 24-ID-C at the Advanced Photon Source, Argonne National Laboratory, funded by the US Department of Energy. We thank L. Jaeger (University of California, Santa Barbara) for discussions and O. Ouerfelli (Memorial Sloan-Kettering Cancer Center, New York) for the synthesis of iridium hexamine. D.J.P. was supported by funds from the National Institutes of Health.

Author Contributions L.H. crystallized the *F. nucleatum* FMN riboswitch and performed binding experiments, A.S. determined the structures and performed footprinting experiments, and A.S. and D.J.P. wrote the manuscript with the assistance of L.H. All authors discussed the results and commented on the manuscript.

Author Information The coordinates of the X-ray structures of the FMN riboswitch are deposited in the RCSB Protein Data Bank under the following accession numbers: bound to FMN, transcribed RNA, 3F2Q; bound to FMN, two-strand composition, 3F4E; bound to riboflavin, 3F4G; bound to roseoflavin, 3F4H; $[Ir(NH_3)_6]^{3+}$-soaked, 3F2T; Cs^+-soaked 3F2X; Ba^{2+}-soaked, 3F2W; Mn^{2+}-soaked, 3F2Y; $[Co(NH_3)_6]^{3+}$-soaked, 3F30. Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to A.S. (serganoa@mskcc.org) or D.J.P. (pateld@mskcc.org).

©2009 Macmillan Publishers Limited. All rights reserved

METHODS

RNA preparation and complex formation. The 112-nucleotide sensing domain of the *F. nucleatum* FMN riboswitch followed by a hammerhead ribozyme was cloned into the pUT7 vector[31] and transcribed *in vitro* using T7 RNA polymerase. The RNA was purified by denaturing polyacrylamide gel electrophoresis and anion-exchange chromatography. Alternatively, the riboswitch was formed by the annealing of two chemically synthesized RNAs (0.4 mM each), 5′-GGAUCUUCGGGGCAGGGUGAAAUUCCCGACCGGUGGUAUAGUCCACGAAAGCUU and 5′-GCUUUGAUUUGGUGAAAUUCCAAAACCGACAGUAGAGUCUGGAUGAGAGAAGAUUC, designed to engineer crystal contacts, in the presence of FMN or its analogues in the binding buffer at 37 °C for 30 min followed by incubation on ice. The 143-nucleotide sensing domain of the *B. subtilis* FMN riboswitch, the 220-nucleotide full-length *F. nucleatum* FMN riboswitch and the 172-nucleotide fragment of the *B. subtilis* lysine riboswitch[27] were transcribed and purified as above. Ligand concentrations were estimated spectrophotometrically using the extinction coefficients $\varepsilon_{450} = 12{,}500\ M^{-1}cm^{-1}$ for FMN, $\varepsilon_{505} = 31{,}000\ M^{-1}cm^{-1}$ for roseoflavin, and $\varepsilon_{445} = 12{,}500\ M^{-1}cm^{-1}$ for riboflavin and lumiflavin.

Crystallization and X-ray crystallography. Crystals were grown at 20 °C using hanging-drop vapour diffusion by mixing 1 μl complex with 1 μl reservoir solution. The FMN-bound complex produces crystals of the saturated yellow colour, characteristic for the oxidized form of FMN, under several conditions. The best crystals were obtained in the solution containing 0.1 M MES-sodium, pH 6.5, 100–200 mM $MgCl_2$ and 6–13% (w/v) PEG 4000 for 1–4 weeks. The crystals of the analogue-bound complexes grew in solution containing 0.1 M Tris-HCl, pH 8.4, 200 mM $MgCl_2$ and 6–10% (w/v) PEG 4000 for about 1 week. The FMN-bound native and heavy-atom-soaked crystals were grown using the transcribed RNA. Although the analogue-bound crystals could grow using the transcribed RNA, resolution of the crystals was improved with the annealed riboswitch.

X-ray diffraction data were reduced using HKL2000 (HKL Research). The structure was determined using 3.0-Å MAD iridium data and autoSHARP[32]. The RNA model was built using TURBO-FRODO (http://www.afmb.univ-mrs.fr/-TURBO-) and refined with REFMAC[33]. The final 2.95-Å riboswitch model contained 109 nucleotides (nucleotides 54–56 were cleaved off during crystallization), one FMN molecule, two potassium and 14 magnesium cations. Mn^{2+}, Cs^{+}, Ba^{2+} and $[Co(NH_3)_6]^{3+}$ cations were positioned based on the anomalous electron density maps (Supplementary Figs 7, 8, 10 and 11). Mg^{2+} and K^{+} cations were modelled according to location of their mimics, coordination geometry and distances[34].

Fluorescence measurements. Fluorescent assays were performed based on intrinsic fluorescence of FMN and its analogues, which become quenched after specific interaction of the ligands with the riboswitch fragment. In all assays, intensity of fluorescence emission was measured at 530 nM with excitation at 450 nM. Each experiment was performed about two to four times at room temperature using a Tecan M1000 fluorimeter.

For the binding affinity measurements, the fragments of the *F. nucleatum* and *B. subtilis* FMN riboswitches were titrated against 6×10^{-8} M FMN or 10^{-8} M analogues. RNA and ligands were premixed in 50 mM Tris-HCl, pH 7.4, 100 mM KCl and 2 mM $MgCl_2$ in the 96 half-area black flat plates for 10–60 min. The *B. subtilis* lysine riboswitch[27] was used as a negative control. After subtraction of the buffer fluorescence and normalization to the free ligand fluorescence, the data were fitted to equation (1):

$$F = 1 + (f-1)\frac{(L_0 + R_0 + K_d) - \sqrt{(L_0 + R_0 + K_d)^2 - 4L_0R_0}}{2L_0} \quad (1)$$

where F is normalized fluorescence intensity, L_0 and R_0 are the concentrations of ligand and RNA, K_d is the apparent dissociation constant, and f is a residual fluorescence intensity at the saturated concentration of ligand, determined by plotting F versus R_0.

For cation-dependence studies, the 112-nucleotide *F. nucleatum* FMN riboswitch was titrated against 6×10^{-8} M of FMN in 50 mM Tris-HCl, pH 7.4, supplemented with 2 mM different cations ($MgCl_2$, $MnCl_2$, $BaCl_2$, $CaCl_2$, $[Ir(NH_3)_6]Cl_3$ and $[Co(NH_3)_6]Cl_3$) in the presence and absence of 100 mM KCl or in the presence of other monovalent cations (NaCl or CsCl). Binding affinities were determined using equation (1). The cation concentration required for the FMN binding was estimated by titration of different cations against a mixture of RNA (2×10^{-7} M) and ligand (6×10^{-8} M) in 50 mM Tris-HCl, pH 7.4. After background subtraction of the fluorescent quenching at each point in the absence of RNA, data were fitted to the Hill equation (2),

$$\theta = [M]^n/(K_d + [M]^n) \quad (2)$$

where θ is the normalized FMN-bound fraction, n is Hill coefficient, $[M]$ is the concentration of cation and K_d is the apparent dissociation constant. As the parameters K_d and n covary, the cation binding was roughly estimated as the ion concentration $[M]_{1/2} = (K_d)^{1/n}$ at which approximately 50% of FMN was bound to RNA.

Footprinting studies. For footprinting experiments[35], the 112-nucleotide *F. nucleatum* riboswitch was radioactively labelled at the 5′ end by the kinase reaction. Samples (20 μl) of the radiolabelled RNA (100,000 c.p.m.) with a final RNA concentration 0.5 μM were preheated at 37 °C for 10 min in 50 mM Na-HEPES, pH 7.9, 50 mM KCl and 2 mM $MgCl_2$. Sixfold excess of FMN was added to the RNA and the mixtures were additionally incubated at 37 °C for 15 min. Cleavage reactions were performed with 0.003 U RNase V1 (Pierce) or 0.25 U RNase T2 (Sigma) at 37 °C for 10 min. Reactions were quenched by the addition of 80 μl cold buffer and were immediately extracted with phenol-chloroform and precipitated by ethanol. Radiolabelled pellets were dissolved and analysed by polyacrylamide gel electrophoresis.

31. Serganov, A. *et al.* Ribosomal protein S15 from *Thermus thermophilus*: cloning, sequencing, overexpression of the gene and RNA-binding properties of the protein. *Eur. J. Biochem.* **246,** 291–300 (1997).
32. de La Fortelle, E. & Bricogne, G. in *Methods in Enzymology* 472–494 (Academic Press, 1997).
33. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D* **53,** 240–255 (1997).
34. Feig, A. L. & Uhlenbeck, O. C. in *The RNA World* 2nd edn (eds Gesteland, R. F., Cech, T. R. & Atkins, J. F.) 287–319 (Cold Spring Harbor Laboratory Press, 1999).
35. Serganov, A., Polonskaia, A., Ehresmann, B., Ehresmann, C. & Patel, D. J. Ribosomal protein S15 represses its own translation via adaptation of an rRNA-like fold within its mRNA. *EMBO J.* **8,** 1898–1908 (2003).

©2009 Macmillan Publishers Limited. All rights reserved

CORRIGENDUM

doi:10.1038/nature07834

Global distribution and conservation of rare and threatened vertebrates

Richard Grenyer, C. David L. Orme, Sarah F. Jackson, Gavin H. Thomas, Richard G. Davies, T. Jonathan Davies, Kate E. Jones, Valerie A. Olson, Robert S. Ridgely, Pamela C. Rasmussen, Tzung-Su Ding, Peter M. Bennett, Tim M. Blackburn, Kevin J. Gaston, John L. Gittleman & Ian P. F. Owens

Nature 444, 93–96 (2006)

165 mammalian species (from a total of 4,269) listed as 'data deficient' by the IUCN were inadvertently omitted from the total and rare species richness values presented. In addition, the spatially corrected *P* values associated with richness correlations presented incorrectly used a one-tailed significance test for positive correlations. Consequently, a number of numerical results in the original paper have changed. Our original conclusions are not affected.

A complete set of revised display items is available for comparison as Supplementary Information to this corrigendum. The range of correlations between total richness values reported in the original paper ($0.79 = r = 0.90$) are unchanged to 2 decimal places, however that between rare species richness values should be $0.22 = r = 0.47$ (originally $0.24 = r = 0.48$). The lower quartile of the range size distribution of mammals should be 41,728 km^2 (originally 41,685 km^2). 19.3% of hotspot cells are common to all three vertebrate classes (originally 17.8%), compared to 2% (originally 2.3%) of rarity hotspots. The median range sizes we reported for illustrative purposes in the Methods section are derived from distribution maps on the analytical grid, not from polygonal areas of occupancy.

We thank Natalie Cooper for discovering the error in the significance testing.

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

©2009 Macmillan Publishers Limited. All rights reserved

ADDENDUM

doi:10.1038/nature07850

Understanding individual human mobility patterns

Marta C. González, César A. Hidalgo & Albert-László Barabási

Nature 453, 779–782 (2008)

The human mobility study relied on an anonymized billing dataset that was previously recorded by a mobile provider as required by law and billing purposes, and not for the purposes of this project. The research was reviewed and approved by the Institutional Review Board (IRB) at the US Office of Naval Research, the main sponsor of the project, and has approval from the Northeastern University's IRB as well, the institution where the work was carried out. As part of the IRB review, the first author, who handled the data, and the PI participated in ethics training sessions at the outset of the study.

©2009 Macmillan Publishers Limited. All rights reserved

CORRIGENDUM

doi:10.1038/nature07881

A burst of segmental duplications in the genome of the African great ape ancestor

Tomas Marques-Bonet, Jeffrey M. Kidd, Mario Ventura, Tina A. Graves, Ze Cheng, LaDeana W. Hillier, Zhaoshi Jiang, Carl Baker, Ray Malfavon-Borja, Lucinda A. Fulton, Lynne V. Nazareth[1], Donna M. Muzny[1], Can Alkan, Gozde Aksay, Santhosh Girirajan, Priscillia Siswara, Lin Chen, Maria Francesca Cardone, Arcadi Navarro, Richard A. Gibbs[1], Elaine R. Mardis, Richard K. Wilson & Evan E. Eichler

[1]Human Genome Sequencing Center, Baylor College of Medicine, One Baylor Plaza, Houston, Texas 77030, USA.

Nature **457**, 877–881 (2009)

In this Letter, Lynne V. Nazareth, Donna M. Muzny and Richard A. Gibbs were inadvertently omitted from the author list.

©2009 Macmillan Publishers Limited. All rights reserved

The digital generation

Next-generation sequencing is pushing gene-expression profiling further into the digital age. But analog methods still have plenty of wind left. Nathan Blow looks at the looming battle over the cell's transcriptome.

Could it be only a matter of time before all gene-expression analysis goes digital? The answer is less straightforward than you might think. When it comes to studying a cell's transcriptome — the collection of all RNA transcripts produced at a specific time — next-generation sequencing promises insights into the way genes are expressed and regulated in cells. Compared with analog methods such as microarrays and real-time PCR, the technology is still in its infancy. But with availability increasing and costs declining, more researchers are turning a curious eye to sequencing's high-throughput digital readouts.

The future of analog techniques — and their possible obsolescence — was one of the questions tackled in a conference panel session held before a packed audience at last year's conference on issues in parallel sequencing at Harvard University Medical School in Boston, Massachusetts. The panel explored the changed landscape of gene-expression analysis since the arrival of next-generation sequencing, and feelings on the subject in both panel and audience were mixed.

"Sequencing allows you to ask many different types of questions and look at transcription from new angles," says Shirley Liu, an associate professor at the Harvard School of Public Health and Dana-Farber/Harvard Cancer Center in Boston, who organized the conference and led the panel. But she thinks sequencing has a long way to go before it reaches the level of adoption that microarrays have among researchers. "My feeling is that if people have never done this kind of work and are trying things for the first time or are asking traditional questions, they are better off with arrays now," she says.

By the numbers

"These days researchers are stepping into a new world when looking at sequencing," says Shawn Baker, a senior product manager for RNA analysis at Illumina in San Diego, California. Although the 'big splash' for most new next-generation sequencing systems comes from deciphering eukaryotic genomes at a breakneck pace, Baker says that digital gene expression and RNA profiling are increasingly popular. He suggests that the company's Genome Analyzer is used for this type of work 30–40% of the time — although he adds that when the major genome centres are subtracted from the user mix that percentage is likely to be much higher.

Four companies now provide next-generation sequencing platforms that support, or plan to support, digital gene-expression applications, but that number is widely predicted to rise in the coming years as developers work on 'next-next-generation' single-molecule sequencing systems. For now, Applied Biosystems (part of Life Technologies) in Foster City, California, recently released a whole transcriptome-expression kit along with a small RNA expression kit for use on its SOLiD 3 platform. Illumina has developed a digital gene-expression kit for RNA analysis and profiling for use on its Genome Analyzer; and 454 Life Sciences, a Roche company located in Branford, Connecticut, uses a serial analysis of gene expression (SAGE) tag-based approach for expression profiling on its GS FLX system.

This level of development and interest among researchers should come as no surprise: gene-expression analysis is an application that seems ideally suited to next-generation sequencing. "For digital gene expression you are just counting the number of times you hit a gene and then assuming that that represents the number of copies of the transcript that you have in your population," says Chad Nusbaum, co-director of the genome sequencing and analysis programme at the Broad Institute and Harvard University, both in Cambridge, Massachusetts. So the more you can count — and next-generation systems can count a lot — the better the measure of copy number for even those rare transcripts in a population.



J. RINN

John Rinn is working on the biological functions of large non-coding RNAs.

"People are sequencing anything from 10 million to 40 million reads in a run," says Baker, "and they are getting a phenomenal level of data." Last year, short-read high-throughput sequencing, or 'RNA-seq', was used to obtain global views of gene expression in human embryonic kidney and B cells[1], to profile transcription in mouse embryonic stem cells[2] and to quantify whole mouse-cell transcriptomes[3].

Those studies also highlight the other benefit that comes with a sequencing-based approach. "Sequencing is the best way to truly profile all aspects of sequence variation," says Roland Wicki, director of Applied Biosystems' SOLiD application development. In addition to looking at RNA expression patterns, RNA-seq can allow researchers to discover new classes of RNA, detect point mutations in expressed transcripts, identify fusion transcripts or uncover new alternative splicing events — discovery applications not possible with other technologies. Using the SOLiD 3 platform, Nicole Cloonan of the Institute for Molecular Bioscience at the University of Queensland in St Lucia, Australia, and her colleagues identified more than 2,000 expressed single nucleotide polymorphisms (SNPs) from embryonic stem cells and determined that the RNA-seq approach could detect 25% more genes than analog microarrays. Ali Mortazavi of the California Institute of Technology, Pasadena, and his



ILLUMINA

Next-generation sequencing platforms such as the Illumina Genome Analyzer are allowing scientists to explore the transcriptome in greater detail than ever before.

© 2009 Macmillan Publishers Limited. All rights reserved



APPLIED BIOSYSTEMS

Applied Biosystems has released a whole-transcriptome-expression analysis kit for the SOLiD 3 platform.

colleagues used Illumina's Genome Analyzer to find 3,500 genes that showed one or more alternative splice forms.

Wicki says that when it comes to the identification of point mutations in transcripts or analysis of allele-specific expression, next-generation sequencing platforms such as the SOLiD 3 system, which uses a unique ligation-based sequencing approach that allows for two-base encoding, tend to be highly accurate. Still, the real value of such systems lies in the ever-increasing numbers of sequencing reads they can generate.

Enriching discovery

Sequencing today is allowing researchers to increase the dynamic range of their investigations simply by increasing the number of sequencing reads analysed. There is a cost to this, though. "High-throughput sequencing of RNA does take a lot of reads to get depth," says John Rinn, an assistant professor at Beth Israel Deaconess Medical Center in Boston, Massachusetts. "So you end up sequencing highly abundant RNAs over and over again before getting to low-abundance stuff." Although this is necessary when it comes to profiling gene-expression patterns, it has led several researchers, including Rinn's group, to develop new methods to enrich for specific populations of RNAs from the transcriptome (see 'Rethinking junk DNA') when it comes to discovery applications.

Several new enrichment procedures focus on eliminating the abundant amounts of ribosomal RNA from a pool of total RNA before the library-generation step of RNA-seq. Invitrogen, part of Life Technologies in Carlsbad, California, recently introduced the RiboMinus kit for RNA-seq applications. The kit depletes ribosomal RNA from a sample by binding the ribosomal RNA to probes containing locked nucleic acids, which are then separated from the sample using magnetic beads.

Another enrichment approach was developed by Evrogen of Moscow, Russia. This uses a duplex-specific nuclease for normalization of RNA transcript levels. Following complementary DNA generation, the templates are denatured and a duplex-specific nuclease is added to the reaction. Although abundant transcripts find matches and become double-stranded, thereby acting as targets for the nuclease, less abundant transcripts take longer to find their partners and so are degraded less frequently.

With or without enrichment for specific RNA populations, the analysis of tens of millions of sequence reads can be a daunting task for most researchers, especially as analysis tools in the digital world are not as advanced as their analog counterparts.

"High-throughput sequencing data analysis is totally different from using arrays," says Liu,

RETHINKING JUNK DNA

When the complete sequence of human chromosome 22 was first published in 1999 (ref. 4), John Rinn, an assistant professor at Beth Israel Deaconess Medical Center and an associate member of the Broad Institute in Cambridge, Massachusetts, got very excited. He was not interested in looking at the map of known protein-coding genes on the chromosome, but rather everything else. "We wanted to see if we could find biologically active molecules in the human genome that no one previously knew about," he says.

Armed with the sequence of an entire chromosome — and a year later the whole human genome — researchers and developers began to create genome-wide tiling microarrays. "By probing these tiling arrays we found out that there are tonnes of biologically active regions by proxy of RNA being made," says Rinn — results he and his colleagues reported in 2003 (ref. 5). Since then, Rinn has focused his efforts on understanding a collection of these RNAs known as large intervening non-coding RNAs (lincRNAs).

"Initially many people thought that this had to be an artefact of the technology: how could there be so many RNA molecules that we have never seen before?" says Rinn. Arguments against a true biological purpose for lincRNAs came largely from the lack of evolutionary conservation within their sequences — conservation implies function, whereas lack of conservation can often imply noise.

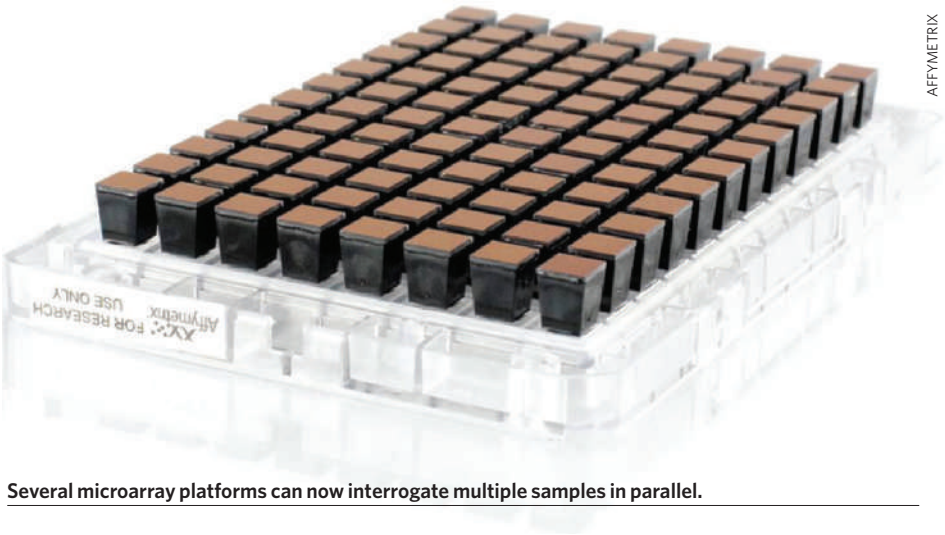
As so few functional lincRNAs had been described, Rinn and his colleagues set out to find more. In 2007 they reported the identification of a new 2.2-kilobase large non-coding RNA, which they called HOTAIR. It played a role in the guiding of chromatin complexes within the cell[6]. Although only a single new functional lincRNA — and still only one of four known to be functional at the time — the discovery gave Rinn an idea on how to enrich for functional lincRNAs from the genome.

"What we did next was to go after things that looked like HOTAIR," he explains. Instead of using an RNA-based approach, the group decided to look at chromatin structure. Histones have clear indications of where active genes start and stop. Using high-throughput chromatin

© 2009 Macmillan Publishers Limited. All rights reserved

adding that at the moment there are few standard tools for analysing the digital gene-expression data sets generated with next-generation sequencing platforms.

"I think in the middle of last year we realized that there was a shift coming from the analog to the digital platform," says Roald Forsberg, director of scientific software solutions at CLC bio in Aarhus, Denmark, which specializes in analysis tools for high-throughput sequencing data sets. This led CLC bio to update its main and genomics software packages to support the analysis of both digital and analog gene-expression data sets. Although Forsberg suspects the shift to digital will take time, he thinks analysis tools with the ability to interrogate both sequencing and microarray data sets will remain critical to researchers for even longer. "There has been much investment in microarrays in both the academic and pharma worlds, using unique tissue samples in a lot of cases, which would be a shame to ignore or just throw away," he says.



AFFYMETRIX

Several microarray platforms can now interrogate multiple samples in parallel.

Higher-level complications

Currently, the biggest challenge for researchers looking at next-generation sequencing approaches is probably the sheer volume of data. "For serial analysis of gene-expression experiments there are some tools, but a lot of them are not well equipped to deal with the amount of tags that come out from high-throughput sequencing," says Liu. This creates the need to perform most data analysis on more powerful computer clusters. "It is not something where you can download a program to run on your laptop," says Liu.

Then there is the challenge of determining where those millions of tags or short sequences come from in the genome. "It is not uncommon for a sequencing platform to generate data and only 40–50% of the data are mappable," says Liu — meaning that the remaining could not be mapped at all or mapped to regions of the genome once considered 'junk'. And for Liu and others the question then becomes, are these biological relevant sequences — for instance, unannotated genes or antisense transcripts — or merely artefacts of the sequencing process?

"I think it is hard to say, but it is probably going to end up being a bit of both," says Jay Shendure, an assistant professor in the department of genome sciences at the University of Washington in Seattle. Shendure says that his group has performed runs on the Illumina Genome Analyzer using genomic DNA where 95% of the reads were mappable, implying that for human DNA libraries with similarly sized read-lengths, the technical artefacts are not so great as to result in only 50% mappable reads. He cautions, however, that there are additional steps in making a sequencing library from RNA that could introduce some artefacts.

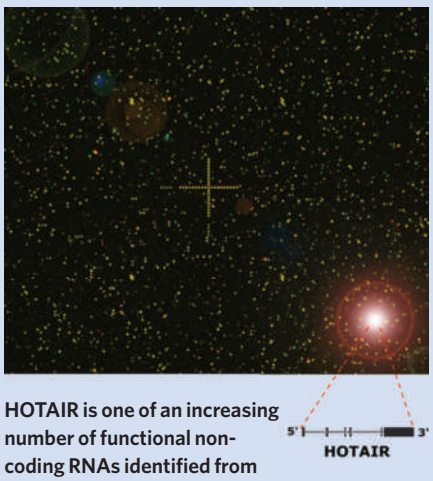
Developers at CLC bio have also experienced problems in mapping short-read RNA-seq data sets. Forsberg hopes that the future use of 'paired-end' reads — in which sequencing information is obtained from both ends of a DNA fragment — might help when mapping back to the genome. He has noticed reservations among researchers when it comes to generating either longer reads or using paired-end protocols for RNA-seq, because this increases the time and cost of sequencing.

Economic advantage

Although next-generation sequencing provides a high-throughput option to look at gene-expression profiles from a small number of samples, it turns out that in the analog world microarrays provide their own high-throughput advantage to researchers. "With microarrays you can profile RNA from a hundred different samples, which would be incredibly expensive to do *de novo* with sequencing," says Rinn.

Lower costs alongside increasing probe density on whole-genome tiling arrays for transcript-mapping applications are keeping microarrays from being lost in the blur of sequencing advances. "I think the array platform is well suited to screening studies for a quick look at the transcriptome or the general clustering of samples," says Baker. He is quick to add that the per-sample cost for microarrays has dropped significantly in recent years, with some genome-wide tiling arrays now costing

immunoprecipitation (ChIP) sequencing on the Illumina Genome Analyzer to look for these marks, Rinn and his colleagues at the Broad Institute developed genome-wide chromatin state maps. Then, just as with his analysis of chromosome 22 almost ten years ago, Rinn says he threw out the known protein-coding genes and looked at what was left. He identified 1,600 other RNAs located by themselves in the middle of nowhere in the genome that look just like HOTAIR[7].



J. RINN

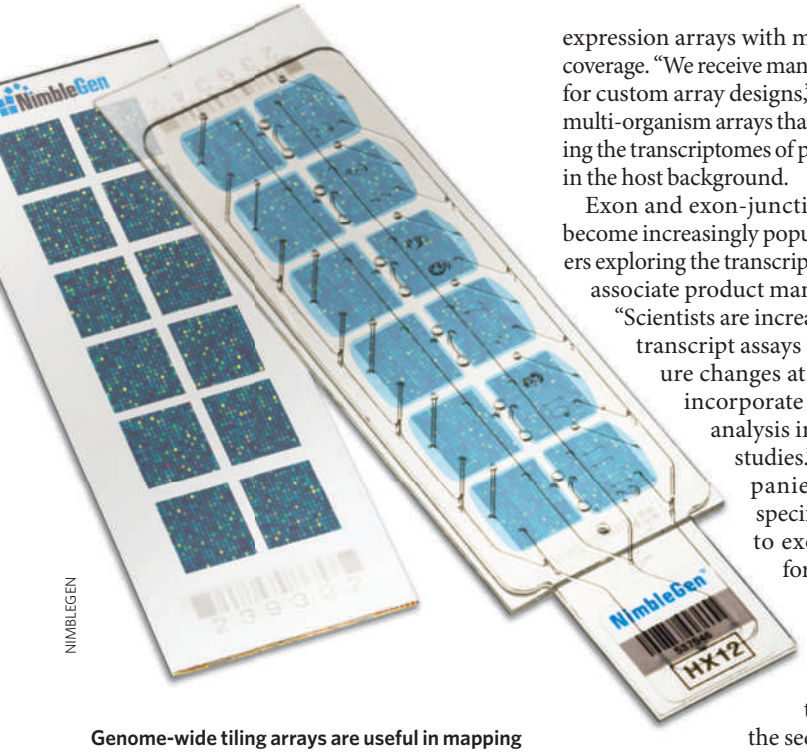
HOTAIR is one of an increasing number of functional non-coding RNAs identified from the human genome.

To determine if some of their newly discovered RNAs were functional, the team took a 'guilt by association' approach, using microarrays to profile a number of the newly identified lincRNAs in 21 different tissue samples while at the same time profiling protein-coding genes in the same tissue samples.

Then they asked the question: which RNAs had similar profiles to protein-coding genes of known function? Their initial analysis was followed by further validation using independent systems. "This has turbo-charged the field, as not only can we identify these things now but we can get a good idea of what they might be doing to test functional relationships," says Rinn.

For Rinn and his colleagues it is now time to muster all the force they can to explore these RNAs. "We are going to throw the Broad kitchen sink at them," says Rinn, who is teaming up with a number of scientific platforms at the Broad Institute to look at the effects of knocking down each newly discovered lincRNA. N.B.

© 2009 Macmillan Publishers Limited. All rights reserved



NIMBLEGEN

Genome-wide tiling arrays are useful in mapping RNAs to specific genomic locations.

less than $100 per sample.

"Resolution is essential for tiling arrays to achieve sensitive and reproducible data, and the more closely spaced the probes are, the more likely they will detect small exons and small RNAs," says Rohaizah James, expression product manager at Roche NimbleGen in Madison, Wisconsin.

High density seems to be the order of the day. Roche NimbleGen now offers a 2.1-million-feature whole-genome tiling array to explore both coding and non-coding parts of the genome. Affymetrix of Santa Clara, California, provides tiling arrays with 6.4 million features consisting of 25 base-pair probes spaced 10 bases apart across the entire human genome for transcript mapping. And Agilent Technologies, also in Santa Clara, California, offers custom tiling arrays with several hundred thousand features.

James says that tiling arrays can be used for quick and accurate empirical annotations of transcriptomes, information that can then be applied to an expanding area of interest among researchers: building custom differential-expression arrays with more comprehensive coverage. "We receive many interesting requests for custom array designs," says James, such as multi-organism arrays that are useful for studying the transcriptomes of pathogenic organisms in the host background.

Exon and exon-junction arrays have also become increasingly popular among researchers exploring the transcriptome, says Joel Fellis, associate product manager at Affymetrix. "Scientists are increasingly using whole-transcript assays and arrays to measure changes at the exon level, and incorporate alternative splicing analysis into their expression studies." A number of companies now offer exon-specific arrays in addition to exon-junction arrays for these studies.

Having been around for some time now, microarrays also tend to offer researchers the security of a standardized technique. "I think that is a very mature tool pack, where people know what to expect," says Liu. The tools to analyse microarrays are more widely available and better developed, making analysis less complicated and easier for even novice researchers. Analog measurements come with a downside, however — one that has hounded microarrays from the start. "There is a frequent lack of inter-platform reproducibility," says Shendure, something that digital, direct-counting platforms could overcome.



J. SHENDURE

Jay Shendure thinks microarrays could migrate towards preparative uses.

Others argue the real challenge for microarrays in the coming years will be to remain up to date. "Our understanding of the transcriptome is constantly evolving, making it difficult for microarrays to stay current," says Wicki. For this reason a number of researchers expect microarrays to migrate towards more targeted applications in the future, perhaps associated with biomarker validation or other diagnostic applications. "I think in terms of basic science research or mechanistic understanding it could be that sequencing will gradually phase out arrays," Liu says.

Although microarrays continue to be the tool of choice for most researchers for gene-expression analysis, looking at copy-number variations within the genome and genotyping, some scientists see even bigger changes for this technology on the horizon. "I think sequencing will penetrate all those markets with increasing force as the cost drops," says Shendure. Although he thinks this will lead to a steady decline in the analytical role of microarrays over time, he strongly suspects their use in preparative applications will expand in the coming years.

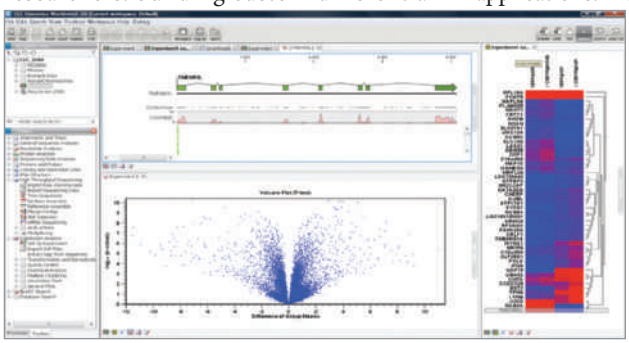
The way ahead

"Arrays will have a role as a super-cheap way of making DNA," says Shendure. At the 2009 Advances in Genome Biology Technology meeting in February at Marco Island in Florida, Shendure set up a workshop on methods to selectively capture subsets of the genome or transcriptome for high-throughput sequencing. Three of the four methods he highlighted relied on microarrays to isolate specific fragments of DNA either by hybridization of genomic libraries directly to the arrays, or by solution-phase hybridization to oligonucleotides cleaved from arrays, or by an approach that converts oligonucleotides released from arrays to inversion probes to capture exons. "We are now able to capture the whole coding genome on a single array," he says.

Developers have recently started supplying arrays for preparative applications as well. Agilent Technologies has partnered with the Broad Institute to provide the SureSelect Target Enrichment System for the Illumina Genome Analyzer, with a version in development for Applied Biosystems' SOLiD platform. The current system uses up to 55,000 biotinylated RNA probes in a single tube to capture particular segments of the genome, but future plans include an array-based partitioning tool for smaller-scale experiments. Last month, Roche NimbleGen introduced a human exome sequence capturing array to prepare samples for sequencing on the GS FLX systems from 454 Life Sciences.

In the end, most researchers think advances in next-generation sequencing for digital gene-expression analysis will continue to push microarrays towards more specific applications. Just how soon researchers leave their analog roots and take that digital plunge, however, will depend on how quickly developers can standardize and support a digital lifestyle. ■

Nathan Blow is technology editor for *Nature* and *Nature Methods*.



CLC BIO

Software programs such as CLC bio's Genomic Workbench are starting to simplify digital gene-expression analysis.

1. Sultan, M. *et al. Science* **321,** 956–960 (2008).
2. Cloonan, N. *et al. Nature Methods* **5,** 613–619 (2008).
3. Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L. & Wold, B. *Nature Methods* **5,** 621–628 (2008).
4. Dunham, I. *et al. Nature* **402,** 489–495 (1999).
5. Rinn, J. L. *et al. Genes Dev.* **17,** 529–540 (2003).
6. Rinn, J. L. *et al. Cell* **129,** 1323–1329 (2007).
7. Guttman, M. *et al. Nature* doi:10.1038/nature07672 (2009).

© 2009 Macmillan Publishers Limited. All rights reserved

COMPANY	PRODUCTS/ACTIVITY	LOCATION	URL
DNA sequencing and analysis			
454 Life Sciences	Automated DNA analysis systems including Genome Sequencer 20 and FLX	Branford, Connecticut	www.454.com
Applied Biosystems	Automated sequencing systems using both Sanger methodology and next-generation SOLiD-based system	Foster City, California	www.appliedbiosystems.com
Beckman Coulter	Automated DNA sequencing systems based on capillary electrophoresis using the Sanger method	Fullerton, California	www.beckmancoulter.com
CLC bio	Developing software platforms for the analysis of next-generation sequencing data sets	Aarhus, Denmark	www.clcbio.com
Cogenics	Contract automated DNA sequencing services using both Sanger-based and next-generation sequencing systems	Morrisville, North Carolina	www.cogenics.com
DNASTAR	Software for DNA sequence and microarray analysis	Madison, Wisconsin	www.dnastar.com
GE Healthcare	Kits for automated DNA sequencing systems	Buckinghamshire, UK	www.gehealthcare.com
Geospiza	Software for management genetic data	Seattle, Washington	www.geospiza.com
Helicos Biosciences	Developing HeliScope Genetic Analyzer based on single-molecule sequencing methodology	Cambridge, Massachusetts	www.helicosbio.com
Illumina	Illumina Genome Analyzer next-generation DNA sequencing system	San Diego, California	www.illumina.com
LI-COR Biosciences	Automated DNA sequencing instrumentation using the Sanger methodology	Lincoln, Nebraska	www.licor.com
MACROGEN USA	DNA sequencing services	Rockville, Maryland	www.macrogenusa.com
Pacific Biosciences	Automated next-generation sequencing instrument	Menlo Park, California	www.pacificbiosciences.com
Promega	*fmol* DNA sequencing system for manual DNA sequencing and promoter mapping applications	Madison, Wisconsin	www.promega.com
RainDance Technologies	Microfluidic-based sequence-enrichment technologies	Lexington, Massachusetts	www.raindancetechnologies.com
Roche Applied Sciences	454 Life Sciences Genome Sequencer 20 and FLX systems	Basel, Switzerland	www.roche.com
Saturn Biotech	Full DNA sequencing service provider	Perth, Australia	www.saturnbiotech.com.au
Sequetech	DNA sequencing services	Mountain View, California	www.sequetech.com
SeqWright	Contract automated DNA sequencing services	Houston, Texas	www.seqwright.com
SoftGenetics	Software analysis tools for genetic research, including for ChIP sequencing and microRNA discovery	State College, Pennsylvania	www.softgenetics.com
DNA microarrays			
Affymetrix	Promoter, ENCODE, chromosome arrays; whole-genome tiling arrays	Santa Clara, California	www.affymetrix.com
Agilent Technologies	CpG islands, promoter, ENCODE arrays; custom tiling arrays; DNA methylation microarray software	Santa Clara, California	www.agilent.com
Expression Analysis	Expression profiling and genotyping services	Durham, North Carolina	www.expressionanalysis.com
Jivan Biologics	Microarrays and PCR for alternative splicing analysis	Larkspur, California	www.jivanbio.com
Marligen Biosciences	Transcription factor analysis, SNP genotyping, DNA and RNA purification kits	Ijamsville, Maryland	www.marligen.com
NanoString Technologies	n-Counter digital gene-expression analysis system using barcodes	Seattle, Washington	www.nanostring.com
NimbleGen	Microarrays, software and services	Madison, Wisconsin	www.nimblegen.com
Oxford Gene Technology	Oligonucleotide arrays	Oxford, UK	www.ogt.co.uk
PerkinElmer	Microarray scanners; genomics-analysis software	Waltham, Massachusetts	www.perkinelmer.com
Phalanx Biotech Group	Whole-genome DNA microarray; gene-expression profiling service	Hsinchu, Taiwan	www.phalanxbiotech.com
Genomic services			
Epigenomics	Biomarker discovery services; kits for methylation analysis of biomarker genes; real-time PCR assays	Berlin, Germany	www.epigenomics.com
Geneservice	Genomic and proteomic resources; contract services for DNA sequencing, microarray analysis and SNP genotyping	Cambridge, UK	www.geneservice.co.uk
Genpathway	Whole-genome ChIP sequencing services; ChIP-on-chip assays	San Diego, California	www.genpathway.com
Miltenyi Biotec	Microarray-based genomic services	Bergisch Gladbach, Germany	www.miltenyibiotec.com
Operon Biotechnologies	Oligonucleotide production; preprinted oligonucleotide arrays and probe sets	Huntsville, Alabama	www.operon.com

© 2009 Macmillan Publishers Limited. All rights reserved

COMPANY	PRODUCTS/ACTIVITY	LOCATION	URL	
QIAGEN	DNA, RNA extraction kits; genotyping; haplotyping	Hilden, Germany	www.qiagen.com	
Sequenom	Genomic analysis services including genotyping, gene expression and methylation analysis, and molecular typing	San Diego, California	www.sequenom.com	
SwitchGear Genomics	Promoter and UTR libraries; functional microarrays; DNA methylation analysis service	Menlo Park, California	www.switchgeargenomics.com	
TeleChem International	Microarray robots, printers and scanners; microarray services	Sunnyvale, California	www.arrayit.com	
Vermillion	Biomarker identification for drug discovery	Fremont, California	www.vermillion.com	

PCR

Abgene	PCR and quantitative PCR reagents	Epsom, UK	www.abgene.com	
Antarus Biotech	PCR instruments	Huntsville, Alabama	www.antarus.net	
Clontech	Variety of molecular biology reagents including PCR products, gene-expression profiling and RNA interference assays	Mountain View, California	www.clontech.com	
Corbett Life Science	Real-time PCR cycler system; other life-sciences instruments	Sydney, Australia	www.corbettlifescience.com	
Finnzymes	PCR reagents and thermal cyclers	Espoo, Finland	www.finnzymes.com	●
SABiosciences	Quantitative real-time PCR assays for ChIP analyses; microRNA arrays; protein phosphorylation assays	Frederick, Maryland	www.sabiosciences.com	
Sigma-Aldrich	Custom oligonucleotides; DNA sequencing systems; RNA interference reagents	St Louis, Missouri	www.sigmaaldrich.com	●
Stratagene	Quantitative PCR reagents and instruments; microarray reagents; RNA and DNA purification kits	La Jolla, California	www.stratagene.com	
TIB Molbiol	Oligonucleotides for PCR, real-time PCR assay development; PCR tests	Berlin, Germany	www.tib-molbiol.com	
Zymo Research	Bisulphite-treated DNA amplification; EZ DNA methylation kit	Orange, California	www.zymoresearch.com	

General

ALEXIS Biochemicals	Reagents for molecular- and cell-biology research	Lausanne, Switzerland	www.alexis-corp.com	
Attagene	Transcription-factor profiling system; software	Research Triangle Park, North Carolina	www.attagene.com	
Cambrex	Molecular- and cell-biology reagents and tools	Charles City, Iowa	www.cambrex.com	
Cole-Parmer	Instruments and reagents	Vernon Hills, Illinois	www.coleparmer.com	
EMD (Novagen)	Reagents and kits for molecular biology	Madison, Wisconsin	www.emdbiosciences.com	
EPICENTRE Biotechnologies	Enzymes for PCR and RT-PCR; DNA and RNA purification and isolation kits	Madison, Wisconsin	www.epibio.com	
Evrogen	Enrichment kits for transcriptomic applications; fluorescent-probe development	Moscow, Russia	www.evrogen.com	
Fisher Scientific	Research equipment, chemicals and kits	Waltham, Massachusetts	www.fishersci.com	●
Hamilton	Robotics for the life sciences	Reno, Nevada	www.hamiltoncomp.com	
Inivtrogen	Kits and reagents for molecular-biology, genomics and cell-biology research	Carlsbad, California	www.invitrogen.com	
Lonza	Molecular-biology reagents and systems	Basel, Switzerland	www.lonza.com	●
Maxim Biotech	PCR reagents and systems; nucleic-acid isolation; custom services including primer design, sample preparation and library construction	South San Francisco, California	www.maximbio.com	
Merck	Chemicals, kits and reagents for molecular and cell biology-related research	Darmstadt, Germany	www.merck.de	
MP Biomedicals	Reagents and chemicals for research	Santa Ana, California	www.mpbio.com	
New England Biolabs	Molecular biology-related reagents, kits and enzymes	Ipswich, Massachusetts	www.neb.com	●
Takara Bio	Reagents, kits and services for genomics and molecular biology research	Shiga, Japan	www.takara-bio.com	●
USB	Chemicals and reagents for molecular biology	Cleveland, Ohio	www.usbweb.com	

● see advertisement

© 2009 Macmillan Publishers Limited. All rights reserved

naturejobs

THE CAREERS MAGAZINE FOR SCIENTISTS

Adjunct and contingent faculty members at universities face trying times. Already fighting for job security, they face an economic downturn that has made their positions even more precarious. These contract workers often have modest salaries and lack benefits. And the proportion of adjuncts versus tenured or tenure-track faculty at US universities — including in the sciences — has grown along with the desire for a flexible and potentially cheaper professorial workforce (see *Nature* **445,** 678–679; 2007). According to a 2007 survey by the US Department of Education, 52% of the higher-education workforce is now either full-time non-tenure-track faculty or part-time/adjunct faculty.

A new US coalition aims to give these staff members a voice. The tentatively named National Coalition of Adjunct Equity formed during a teleconference of 14 adjuncts at US universities. The recession helped trigger the organization's inception, according to co-chair Maria Maisto, an associate lecturer in English composition at the University of Akron, Ohio.

Could the new body help? It couldn't hurt. Other US organizations already fight for adjunct rights, such as the American Federation of Teachers, which includes adjuncts in its labour union, and the American Association of University Professors, which is drafting guidelines for universities on how they can convert part-time faculty members to full time.

But these bodies are not devoted solely to adjunct concerns. "Interests of tenure-track people are not always in line with non-tenure-track," says Maisto, adding that some states make it difficult to unionize. "We wanted to put together an organization with no official position on unions."

But there's even more at stake. When employees lack security or influence, they tend to take fewer risks and not challenge the status quo. More adjuncts means more professors unwilling to present potentially innovative ideas to department heads, a tendency detrimental to institutions and students. A new organization that can help counter this trend would indeed be worthwhile.

Gene Russo is editor of *Naturejobs*.

CONTACTS
Editor: Gene Russo
Assistant editor: Karen Kaplan
e-mail: naturejobseditor@naturedc.com

European Head Office, London
The Macmillan Building,
4 Crinan Street, London N1 9XW, UK
Tel: +44 (0) 20 7843 4961
Fax: +44 (0) 20 7843 4996
e-mail: naturejobs@nature.com

European Sales Manager:
Dan Churchward (4966)
e-mail: d.churchward@nature.com
Assistant European Manager:
Nils Moeller (4953)
***Nature*events:**
Ghizlaine Ababou (+44 (0) 20 7014 4015)
e-mail: g.ababou@nature.com
Southwest UK/RoW:
Alexander Ranken (4944)

Northeast UK/Ireland:
Matthew Ward (+44 (0) 20 7014 4059)
France/Switzerland/Belgium:
Muriel Lestringuez (4994)
Scandinavia/Spain/Portugal/Italy:
Evelina Rubio-Hakansson (4973)
North Germany/The Netherlands/Eastern Europe: Kerstin Vincze (4970)
South Germany/Austria:
Hildi Rowland (+44 (0) 20 7014 4084)
Advertising Production Manager:
Stephen Russell
To send materials use London address above.
Tel: +44 (0) 20 7843 4816
Fax: +44 (0) 20 7843 4996
e-mail: naturejobs@nature.com
***Naturejobs* web development:** Tom Hancock
***Naturejobs* online production:** Dennis Chu

US Head Office, New York
75 Varick Street, 9th Floor,
New York, NY 10013-1917
Tel: +1 800 989 7718
Fax: +1 800 989 7103
e-mail: naturejobs@natureny.com
US Sales Manager: Ken Finnegan

India
Vikas Chawla (+91 1242881057)
e-mail: v.chawla@nature.com

Japan Head Office, Tokyo
Chiyoda Building, 2-37 Ichigayatamachi,
Shinjuku-ku, Tokyo 162-0843
Tel: +81 3 3267 8751
Fax: +81 3 3267 8746

Asia-Pacific Sales Manager:
Ayako Watanabe (+81 3 3267 8765)
e-mail: a.watanabe@natureasia.com
Business Development Manager, Greater China/Singapore:
Gloria To (+852 2811 7191)
e-mail: g.to@natureasia.com

© 2009 Macmillan Publishers Limited. All rights reserved

NATURE|Vol 458|12 March 2009



Tori Barron used community college to make a transition to a biotech career.

UNIV. WISCONSON-MADISON

BIOTECH, THE COMMUNITY WAY

Community-college programmes in the United States can be a quick, cheap route into the pharmaceutical industry. **Ted Agres** talks to scientists making the grade.

Researchers at US drug and biotechnology companies are nervous about their job prospects. So are students hoping to work in industry. But the dozen or so students about to graduate with an associate's degree from Mount Wachusett Community College in Devens, Massachusetts, are likely to have their pick of well-paying jobs in biomanufacturing — across the street, at Bristol-Myers Squibb's new US$750-million cell-culture facility.

"These graduates will be highly sought after," predicts Richard Elmer, the company's manager of learning and development. "There's going to be a talent war with other industry in the local area."

That's good news for the graduates of the two-year course, especially with US biotech companies shedding jobs and research institute endowments falling (see 'Good bet'). "These students will be able to jump into jobs from day one," says Lara Dowland, chair of the new biotechnology and biomanufacturing programme at Mount Wachusett, near Boston.

There are about 1,200 US community and technical colleges granting two-year degrees. They enrol nearly 12 million students, almost half of all undergraduates in the United States. Once seen as learning centres for underperforming students hoping to get any career foothold, these bargain-priced schools have become popular as fees at four-year colleges soar. Fledgling scientists are among those who have taken interest.

Symbiosis at work

The relationship between Bristol-Myers Squibb and Mount Wachusett was crafted from the ground up two years ago — the planning even covered the location of buildings. The company provided specialists in good manufacturing practices and other areas so the school could shape its curriculum to match the firm's needs. The result is a range of courses on biotech basics, the use and calibration of standard lab instruments, hands-on production of human proteins from bacteria, mammalian and yeast cells, and protocols for DNA isolation, gene mapping, expression and regulation.



Genentech's Hank Stern.

The company offered advice on layout, and donated computers, meters, scales and high-performance liquid chromatographs. On top of that, Mount Wachusett last year won a $1.6-million grant from the Department of Labor to finish outfitting the labs with state-of-the-art equipment, provide scholarships and hire training faculty. "It's really a win–win situation," Elmer says.

Bristol-Myers Squibb has hired some 120 full-time employees and expects to fill 350 positions, including 200 for community-college graduates, by the time the cell-culture facility begins commercial production in 2011. "We are qualifying, validating and commissioning equipment right now," Elmer says. "The new graduates will have practical, hands-on experience so they can quickly transition into our labs."

The collaboration between Bristol-Myers Squibb and Mount Wachusett is typical of a growing number of symbiotic relationships between community colleges and biotech and biomanufacturing companies. From 1998 to 2007, the number of active biotech and pharmaceutical programmes at two-year colleges soared from 49 to more than 90. Nearly all are tailored to meet the needs of nearby companies.

"These programmes don't exist at every community college, and they probably shouldn't. They are very specific and require lots of equipment," says Elaine Johnson, director of Bio-Link, a national educational centre based at the City College of San Francisco that

© 2009 Macmillan Publishers Limited. All rights reserved

supports high-tech employment through a grant from the National Science Foundation.

In addition to offering core classes in biology, chemistry, mathematics and communications, US community and technical colleges provide education and training customized to their region's specialist areas. That includes biomanufacturing in the northeast, agricultural bioprocessing in the midwest, medical devices and bioinformatics in the Pacific northwest, and clinical trials and regulatory affairs in the southeast. Stem-cell programmes are also prevalent in California and Wisconsin, where state support or university research is strong.

At an average age of 29, community-college students tend to have more experience than their four-year counterparts. "There is a common misperception that community colleges serve mostly students right out of high school," Johnson says. "Although we do serve those students, there are also a huge number of people seeking new skills, including those with bachelor's and master's degrees seeking to change careers."

Tori Barron, 42, with a degree in psychology, decided to re-enter the workforce after staying at home for seven years to see her three children off to pre-school. In 2002 she enrolled in a post-baccalaureate certificate programme in biotech at Madison Area Technical College in Wisconsin. "The college was very much a springboard for my entry to the biotech workforce," Barron says. She went on to earn a master's degree in biotech from the University of Wisconsin–Madison in 2005. She is now a senior research specialist at the WiCell Research Institute, working on comparative genomic hybridization of human embryonic stem cells. The community-college approach, she says, "is a bargain and offers hands-on training in a cutting-edge field by outstanding faculty".

Holders of two-year degrees as an associate of science or of applied science have a range of options, including microbiologist and analytical chemistry lab technician, specialist in instrumentation and calibration, and posts in validation and quality assurance or environmental health and safety. Graduates can also specialize in areas such as marketing, customer relations, and legal and regulatory compliance.

"Technicians are doing things in labs today that only PhDs did 15 to 20 years ago," says Russ Read, executive director of the National Center for the Biotechnology Workforce at Forsyth Technical Community College in Winston–Salem, North Carolina.

Hank Stern, associate director for manufacturing collaborations at Genentech in South San Francisco, agrees. "These are not trivial jobs by any means," he says. At his company, community-college-trained technicians are operating 25,000-litre cell-culture fermenters producing drugs worth tens of millions of dollars. "A single mistake has a huge impact both in terms of the product we supply to our patients and to the finances of the company," Stern says.



D. ARMSTRONG/MOUNT WACHUSETT COMMUNITY COLLEGE

Class act: Bristol-Myers Squibb's Richard Elmer and local college educator Lara Dowland.

Never too late

Most biotech programmes encourage students to do an internship at a local life-sciences company or lab. Jim Crawford, a graduate of Forsyth Tech, took an unpaid internship at Anthony Atala's lab at the Wake Forest Institute for Regenerative Medicine in Winston–Salem. After the internship, he was offered a full-time job expanding and differentiating stem cells. Two months later, he became a tissue- and cell-culture technologist, managing five rooms used by 90 researchers and with purchasing responsibility for more than $1.5 million in materials annually. Crawford, 58, went to community college after being laid off three times from the textile and electronics industries. "It was the perfect vehicle to both update my credentials and allow me to pursue my goals sought over 30 years ago," he says.

His associate's degree let him start contributing at once. "I was able to walk in and start culturing cells. I was training PhDs and postdocs who had never had that experience," Crawford says. Management skills acquired over 30 years make him well suited for his next step: managing Wake Forest's new manufacturing facility once it is constructed. "It's a big honour and I owe it to the associate's degree programme," he says.

Community colleges also offer short courses and certificate programmes in areas such as immunoassay and Western blotting — and provide networking opportunities. "This is often how students get a toe in the door, which is probably the biggest challenge in finding a job," says Jo-Anne Hongo, a biotech instructor at City College of San Francisco and scientific research manager in the department of antibody engineering at Genentech. "I write a lot of letters of reference," Hongo says. "It's extremely rewarding to see our students get jobs in this highly competitive market." ■

Ted Agres is a freelance writer based in Laurel, Maryland.

GOOD BET

Biomanufacturing is one field that continues to have jobs potential, despite the downturn. "A company manufacturing a drug that has taken 15 years and $1 billion to get approved is not going to stop operations simply because the economy gets a little weak," says Sonia Wallman, director of the Northeast Biomanufacturing Center and Collaborative, a group funded by the US National Science Foundation that works with drug and biomanufacturing companies in the region.

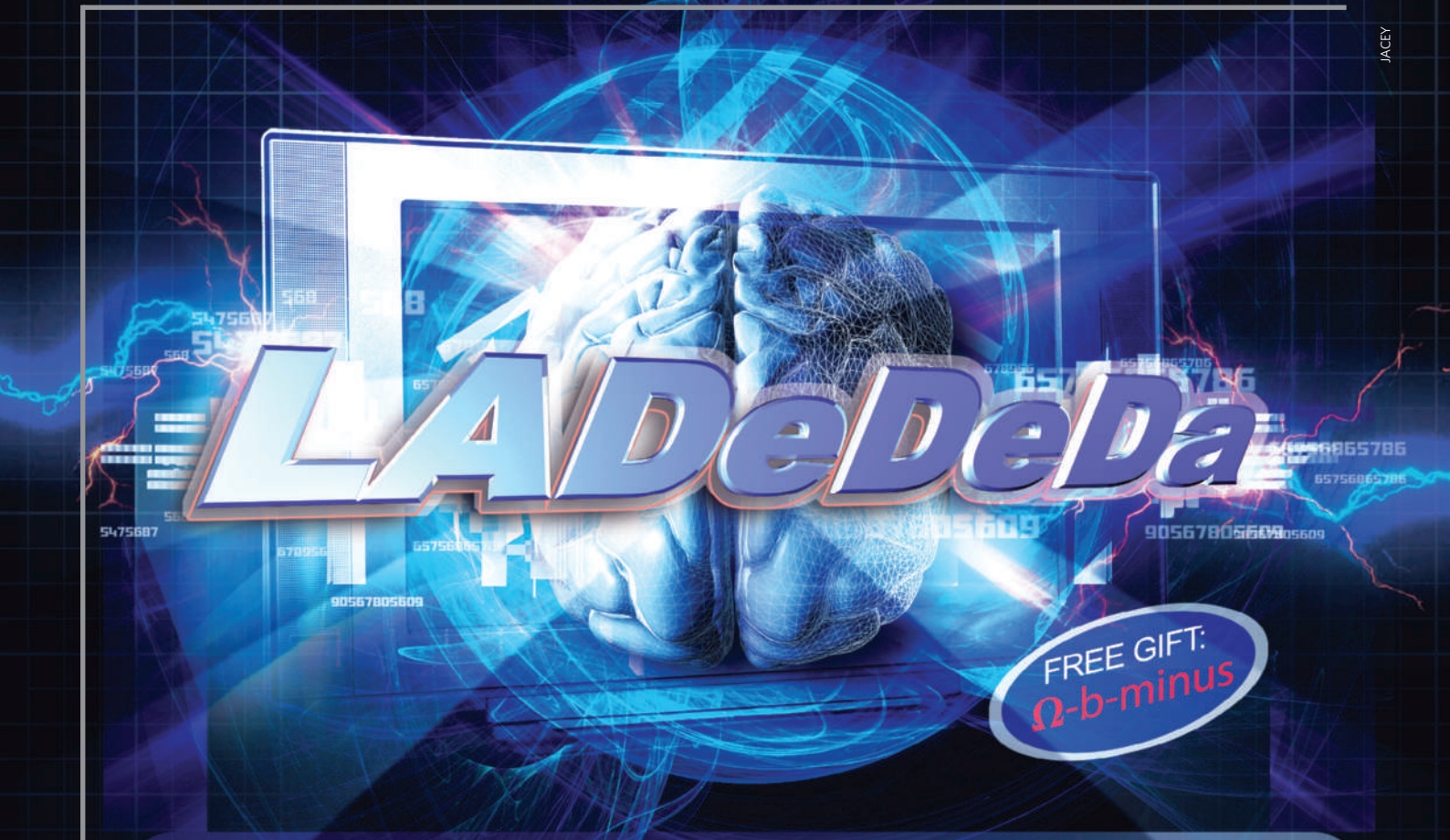
Salaries for biomanufacturing technicians are 28% higher than those in other manufacturing industries. They range from the mid-$30,000s to nearly $60,000, depending on the region, availability of jobs, and whether the employer is a university, research institute or for-profit company. This isn't bad compared with many US postdocs, who often earn less than $45,000 in their first year at work.

About 14,000 biological and chemical technicians were employed at more than 2,500 drug and medicine manufacturing facilities around the country in 2006 (the most recent year available). They represent nearly 5% of the biomanufacturing workforce.

Demand for these positions is expected to grow 26% by 2016, according to the US Bureau of Labor Statistics — although that projection was made before the current global economic downturn. **T.A.**

© 2009 Macmillan Publishers Limited. All rights reserved

JACEY



LADeDeDa Enhancement Kit Includes :

- **RithChek™**
- Waxahachie commemorative bumper sticker
- Mini Hadron Collector
- Slide Rule

Only €20,00 or USD $76.30 !

FUTURES

It is a pleasure to announce the first annual **LADeDeDa**.

Spell-checkers and grammar-checkers have rendered obsolete the quaint fetishes of schoolteachers and dictionaries, liberating writers from the crippling fear of not knowing the difference between its and it's, or there and their, since spell-check doesn't know the difference either. And by preventing the perpetration of any sentence longer than ten words, grammar-check has courageously freed us, at last, from the antiquated convolutions of syntax. Between them, these programs have obviated the need for training in the use of the language. Impossible to imagine that it was once necessary to endure literally years of rigorous study, beginning at the age of six or even younger, and ending only after a decade or more, in order to be able to write English! So gruelling was the training that for centuries it was generally forbidden to girls, lest it interfere with their reproductive capacity. Now, of course, all that drudgery and danger has been done away with. Anyone who can learn to use a letter keyboard has all the freedom of expression he or she can desire, and is free at once to write a masterpiece.

Doomsayers and elitists have proclaimed that the problem of innumeracy may be even greater than that of illiteracy, but all that is behind us too. Just as the pocket calculator freed us from the horrors of the times-tables and the slide rule, now, with **RithChek™** installed in every computer, the whole problem of mathematics vanishes. Nothing is required but moderate keyboard proficiency (thumbs will do), the ability to count, and familiarity with a few elementary symbols, such as +, −, ×, (), √−1.

Any calculation, the simplest addition or the most complex computation, from two times two to the billionth place of π, can be instantly checked for accuracy and corrected. Using **RithChek™**, powerful but intuitive, millions of people who have been unfairly barred from achievements in theoretical and applied science by a mere lack of interest, ability, or training in mathematics are freed to turn their minds loose on the great problems of physics, astrophysics, cosmology, and engineering.

As a natural development, this week the NGnius Coprolation announced its sponsorship of the first **Linear Accelerator Design Development Day**, to run from 00:00:01 to 24:00:00 UTC, 31 August 2009.

LADeDeDa is not a competition, and no winners will be selected. But the directors of cyclotrons, superconducting super colliders, and other linear accelerators at Berkeley, CERN, Brookhaven, Fermilab, Oakridge and elsewhere should make ready for the thousands of great new plans and projects that will be pouring in on them at the close of the event.

A middle-school student might revolutionize the principles of the particle accelerator; a soccer mom might crack the secret of the Higgs boson. In the sciences as in literature, the playing field is level now. All any one of us needs to be a Tolstoy or an Einstein is a belief — a dream — and a laptop. ■

Watch for the gut-wrenching, spin-changing quantum entanglement of Ursula K. Le Guin and Vonda N. McIntyre in the brilliant RithCheck™ infomercial.
www.ursulakleguin.com
www.vondanmcintyre.com

© 2009 Macmillan Publishers Limited. All rights reserved